

**DISERTASI**

**DETEKSI TERSANGKA DENGAN PENGENALAN WAJAH  
BERBASIS VISI KOMPUTER**

***FACE RECOGNITION FOR SUSPECT DETECTION USING  
COMPUTER VISION***

**MUHAMMAD ARAFAH  
D053171006**



**PROGRAM STUDI S3 TEKNIK ELEKTRO  
FAKULTAS TEKNIK  
UNIVERSITAS HASANUDDIN  
MAKASSAR  
2022**

**DETEKSI TERSANGKA DENGAN PENGENALAN WAJAH  
BERBASIS VISI KOMPUTER**

**DISERTASI**

Sebagai Salah Satu Syarat Untuk Mencapai Gelar Doktor

Program Studi  
Doktor Teknik Elektro

Disusun dan Diajukan oleh

**MUHAMMAD ARAFAH  
D053171006**

Kepada

**PROGRAM STUDI S3 TEKNIK ELEKTRO  
FAKULTAS TEKNIK  
UNIVERSITAS HASANUDDIN  
MAKASSAR  
2022**

**LEMBAR PENGESAHAN DISERTASI**

**DETEKSI TERSANGKA DENGAN PENGENALAN WAJAH  
BERBASIS VISI KOMPUTER**

Disusun dan diajukan oleh

**MUHAMMAD ARAFAH**

**D053171006**

Telah dipertahankan di hadapan Panitia Ujian yang dibentuk dalam rangka Penyelesaian Studi Program Doktor Program Studi Teknik Elektro Fakultas Teknik Universitas Hasanuddin pada tanggal 27 April 2022 dan dinyatakan telah memenuhi syarat kelulusan.

Menyetujui  
Promotor,

Prof. Dr. Ir. H. Andani Achmad, M.T.  
NIP. 19601231 198703 1 002

Co. Promotor,

Dr. Indrabayu, S.T., M.T., M.Bus. Sys  
NIP. 19750716 200212 1 004

Ketua Program Studi,

Prof. Dr. Ir. H. Andani Achmad, M.T.  
NIP. 19601231 198703 1 022

Co. Promotor,

Dr. Eng. Ir. Intan Sari Areni, S.T., M.T.  
NIP. 19750203 200012 2 002

Dekan Fakultas Teknik,

Prof. Dr. Eng. Ir. Muhammad Isran Ramli, S.T., M.T.  
NIP. 19730926 200012 1 002

## PERNYATAAN KEASLIAN DISERTASI

Yang bertanda tangan di bawah ini:

Nama : Muhammad Arafah  
Nomor Mahasiswa : D503171006  
Program Studi : S3 Teknik Elektro

Menyatakan dengan sebenarnya bahwa Disertasi yang saya tulis benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambilan tulisan atau pemikiran orang lain. Apabila dikemudian hari terbukti atau dapat dibuktikan bahwa sebagian atau keseluruhan Disertasi ini hasil karya orang lain, saya bersedia menerima sanksi atas perbuatan tersebut.

Makassar, 27 April 2022

Yang Menyatakan,



Muhammad Arafah

## **PRAKATA**

Puji syukur senantiasa penulis panjatkan, atas keridhaan Allah SWT yang telah melimpahkan rahmat dan rahim-Nya serta hidayah-Nya, sehingga penyusunan laporan disertasi ini dapat terselesaikan.

Pada penyusunan disertasi ini terdapat banyak kendala yang penulis alami, dan tidak akan selesai tanpa bantuan dari banyak pihak, baik yang sifatnya moril maupun materil. Oleh karena itu pada kesempatan ini, penulis menyampaikan ucapan terima kasih banyak yang sedalam-dalamnya serta penghargaan yang maha tinggi kepada:

1. Bapak Prof. Dr. Ir. H. Andani Achmad, M.T., sebagai ketua komisi penasehat (promotor) sekaligus sebagai penilai disertasi, Bapak Dr. Indrabayu, S.T., M.T., M. Bus.Sys., dan Ibu Dr. Eng. Ir. Intan Sari Areni, S.T., M.T., sebagai anggota komisi penasehat (kopromotor) sekaligus sebagai penilai disertasi, atas motivasi dan bimbingan yang diberikan sehingga penyusunan disertasi ini dapat diselesaikan. Melalui perantaraan beliau bertiga dengan keikhlasannya yang dibarengi kesabaran dan perhatiannya, telah memberikan pencerahan berupa saran maupun koreksi baik pada aspek penulisan, metodologi serta penyajian isi dari disertasi ini secara keseluruhan.
2. Bapak Dr. Ir. H. Rhiza S Sadjad, MSEE., Bapak Dr. Adnan, S.T., M.T., Ibu Dr. Ejah Umraeni Salam, S.T., M.T., dan Ibu Dr. Eng. Ir. Dewiani, M.T., sebagai komisi penguji yang telah memberikan saran dan masukan guna perbaikan penulisan disertasi ini.

3. Rektor Universitas Hasanuddin, Bapak Prof. Dr. Ir. Jamaluddin Jompa, M.Sc., Dekan Sekolah Pascasarjana Universitas Hasanuddin, Bapak Prof. dr. Budu, Ph.D., Sp.M(K)., M.Med.Ed., Dekan Fakultas Teknik, Bapak Prof. Dr. Eng. Ir. Muhammad Isran Ramli, S.T., M.T., Ketua Departemen Teknik Elektro, Ibu Dr. Eng. Ir. Dewiani, M.T., Ketua Program Studi S3 Teknik Elektro, Bapak Prof. Dr. Ir. H. Andani Achmad, M.T., dan seluruh dosen Teknik Elektro Universitas Hasanuddin yang telah memberi kesempatan dan bekal kepada penulis berupa ilmu dan wawasan yang sangat berarti dan memiliki pengaruh yang signifikan pada penyelesaian disertasi ini. Selanjutnya kepada seluruh staf Teknik Elektro yang telah memberikan layanan pada aspek administrasi selama penulis menuntut ilmu di Universitas Hasanuddin.
4. Ibu Alm. Marwah Made Ali, S.E., Ketua Yayasan Pendidikan Kartini Ujung Pandang (Yapenka), Ibu Asnimar, S. Kom., sebagai sekretaris Yapenka dan Ibu Ramlah P, S.E., M.M., selaku bendahara Yapenka yang telah banyak memberikan bantuan, baik secara moril maupun materil, dalam hal ini memberikan bantuan pendanaan selama penulis menjadi mahasiswa di program studi S3 Teknik Elektro Universitas Hasanuddin.
5. Bapak Dr. Askar Taliang, M.Si., Selaku Rektor Universitas Teknologi Akba Makassar (Unitama) yang telah memberikan arahan dan ijin belajar dalam bentuk rekomendasi kepada penulis untuk melanjutkan pendidikan di Universitas Hasanuddin.

6. Rekan kerja di unsur pimpinan Unitama dalam hal ini Wakil Rektor I, Wakil Rektor II dan Wakil Rektor III. Kemudian Ketua Prodi Sistem Informasi, Ketua Prodi Teknik Informatika, Ketua Prodi Bisnis Digital, Ketua Prodi Teknologi Informasi, Ketua Prodi Pendidikan Teknologi Informasi, Kepala LP3M, Kepala BAAK, kepala BAUK, Kepala Perpustakaan, Kepala UPT ICT dan kepala UPT Bahasa. Selanjutnya kepada teman-teman dosen dan staf di Unitama atas pengertiannya selama masa studi penulis.
7. Teman-teman di grup riset *Artificial Intelligence and Multimedia Processing* (AIMP) dan teman seperjuangan mahasiswa S3 program studi Teknik Elektro, khususnya angkatan 2017 atas *sharing* dan bantuannya dalam penyelesaian disertasi ini.

Selanjutnya penulis menghaturkan ucapan terima kasih dan penghargaan yang setinggi-tingginya kepada kedua orang tua tercinta. Ayahanda Alm H Bade dan Hj Saidah yang telah merawat, membesarkan dan mendidik serta senantiasa mendoakan ananda agar menjadi hamba Allah SWT yang senantiasa dimudahkan segala urusannya dan selamat di dunia serta di akhirat kelak. Juga kepada semua saudara, kakak dan adik serta ipar, terkhusus kepada kakak penulis Alm Muhammad Yusuf, yang menghembuskan nafas terakhirnya di malam tahun baru 2021 atas doa, motivasi dan bantuannya serta dukungannya begitu pula kepada seluruh keluarga besar penulis.

Ucapan terima kasih dan penghargaan terkhusus kepada istri tersayang Tatik Maslihatin, S.T., M. Kom., atas segala pengorbanan, ketulusan, pengertian, kesabaran, motivasi dan keikhlasannya dalam mendampingi penulis. Begitu juga kepada anak-anakku Fatimah Az-Zahra (Farah), Muhammad Husain Al-Fatih (Uceng), Muhammad Mushab Al-Ayyubi (Ayyub), Alifah Al-Husna (Ifah), Afifah Al-Izzah (Izzah) dan Muhammad Alif Al-Faruq (Alif) atas kesabaran, pengertian dan keikhlasannya dalam memberikan semangat maupun motivasi selama penulis menempuh program pendidikan ini.

Akhirnya, pada penyusunan disertasi ini, penulis telah berusaha secara maksimal untuk menyajikannya secara optimal, namun tidak menutup kemungkinan di dalamnya masih terdapat banyak kekurangan. Olehnya itu pada kesempatan ini penulis dengan kerendahan hati mengharapkan masukan atau saran yang sifatnya membangun demi kesempurnaan penyajian disertasi ini.

Makassar, 27 April 2022

Muhammad Arafah

## ABSTRAK

MUHAMMAD ARAFAH, Deteksi Tersangka dengan Pengenalan Wajah Berbasis Visi Komputer (dibimbing oleh Andani Achmad, Indrabayu dan Intan Sari Areni).

Penelitian ini bertujuan untuk mengenali wajah tersangka yang melewati fasilitas umum dengan berbagai kondisi. Penempatan kamera *Closed Circuit Television* (CCTV) dengan jarak yang berbeda-beda, penggunaan oklusi dengan mengenakan kerudung dan pemanfaatan gambar dengan resolusi yang berbeda-beda termasuk resolusi rendah. Pengambilan data uji diskenariokan pada area pemeriksaan keberangkatan penumpang di bandara dengan menggunakan kamera CCTV, untuk data latih menggunakan kamera *Digital Single Lens Reflex* (DSLR). Terdapat dua tahapan yang digunakan yakni deteksi wajah dan pengenalan wajah, untuk tahapan deteksi wajah menggunakan metode *Viola-Jones*, sedangkan untuk pengenalan wajah menggunakan metode *Histogram of Oriented Gradients* dan *Multiclass Support Vector Machine* (HOG-MSVM) dengan menggunakan *pre-processing* (*Brightness Enhancement*, *Contrast Limited Adaptive Histogram Equalization* (CLAHE), *Grayscale*), selanjutnya metode *Convolutional Neural Network* dan *Cosine Similarity* (CNN-CS). Penerapan metode tersebut digunakan pada tiga kasus, kasus pertama adalah pengenalan wajah dengan kondisi penempatan kamera CCTV dengan jarak yang berbeda-beda dengan hasil terbaik pada jarak 300 cm dengan akurasi 86,77% (HOG-MSVM) dan 97,20% (CNN-CS). Kasus kedua yakni pengenalan wajah menggunakan oklusi dengan mengenakan kerudung, terdapat empat skenario yakni data latih tanpa kerudung untuk data uji tanpa kerudung dan dengan kerudung, kemudian data latih dengan kerudung untuk data uji tanpa kerudung dan dengan kerudung. Hasil terbaik pada kasus ini adalah pada skenario data latih tanpa kerudung dengan data uji tanpa kerudung dengan hasil akurasi 85% (HOG-MSVM) dan 99,68% (CNN-CS). Kasus ketiga adalah pemanfaatan gambar dengan resolusi yang berbeda-beda termasuk resolusi rendah menggunakan metode CNN dengan arsitektur ResNet50, ArcFace sebagai loss function pada proses training dan Cosine Similarity untuk proses identifikasi wajah. ResNet50 dan ArcFace menggunakan ukuran embedding sebesar 512 dan pada proses trainingnya, parameter scale dan margin pada ArcFace adalah 64 dan 0,5. Hasil dari metode CNN-CS menunjukkan bahwa meskipun resolusi gambar terkategori rendah sistem dapat mengenali dengan baik, yakni menghasilkan akurasi sebesar 99,35% untuk gambar dengan resolusi 64 piksel dan 97,90% untuk gambar dengan resolusi 32 piksel.

**Kata kunci:** Deteksi Wajah, Pengenalan Wajah, Tersangka, *Brightness Enhancement*, *Contrast Limited Adaptive Histogram Equalization*, *Grayscale*, *Histogram of Oriented Gradients*, *Multi-class Support Vector Machine*, *Convolutional Neural Network*, *Cosine Similarity*.

## ABSTRACT

MUHAMMAD ARAFAH, Face Recognition for Suspect Detection Using Computer Vision (supervised by Andani Achmad, Indrabayu, and Intan Sari Areni).

The aim of this research is to identify the faces of suspects who pass through public facilities with various conditions. The Closed-Circuit Television (CCTV) cameras are placed with different distances, the use of occlusion by wearing a veil and the use of images with different resolutions were included with low resolution. The scenario test data were obtained in the passenger departure inspection area at the airport using CCTV cameras. Meanwhile, for the training data used a Digital Single Lens Reflex (DSLR) camera. There are two stages used, namely face detection and face recognition. The face detection stage used the Viola-Jones method, while the face recognition used the Histogram of Oriented Gradients method and Multiclass Support Vector Machine (HOG-MSVM) using pre-processing (Brightness Enhancement, Contrast Limited Adaptive Histogram Equalization - CLAHE, Grayscale) and Convolutional Neural Network and Cosine Similarity (CNN-CS) methods. The application of this method was used in three cases, the first case was face recognition with CCTV camera placed at different distances with the best results at a distance of 300 cm with 86.77% accuracy for HOG-MSVM and 97.20% for CNN-CS. The second case was facial recognition using occlusion by wearing a veil. There were four scenarios such as training data without a veil for the test data with a veil and without a veil, then training data with a veil for test data with a veil and without a veil. The best result in this case was the scenario of training data without veil with test data without veil with an accuracy of 85% for HOG-MSVM and 99.68% for CNN-CS. The third case was the use of images with different resolutions including low resolution uses CNN method with ResNet50 architecture, ArcFace as a loss function in the training process and Cosine Similarity for the face identification process. ResNet50 and ArcFace use an embedding size of 512 and in the training process, ArcFace's scale and margin parameters are 64 and 0,5. The results from the CNN-CS method showed that the image resolution was categorized as low, the system can recognize it well, which produces an accuracy of 99.35% for images with a resolution of 64 pixels and 97.90% for an image with a resolution of 32 pixels.

**Keywords:** Face Detection, Face Recognition, Suspect, Brightness Enhancement, Contrast Limited Adaptive Histogram Equalization, Grayscale, Histogram of Oriented Gradients, Multi-class Support Vector Machine, Convolutional Neural Network, Cosine Similarity.

## DAFTAR ISI

|   |             |
|---|-------------|
| <b>HALAMAN JUDUL</b> .....                                | <b>i</b>    |
| <b>LEMBAR PENGESAHAN DISERTASI</b> .....                  | <b>ii</b>   |
| <b>PERNYATAAN KEASLIAN DISERTASI</b> .....                | <b>iii</b>  |
| <b>PRAKATA</b> .....                                      | <b>iv</b>   |
| <b>ABSTRAK</b> .....                                      | <b>viii</b> |
| <b>ABSTRACT</b> .....                                     | <b>ix</b>   |
| <b>DAFTAR ISI</b> .....                                   | <b>x</b>    |
| <b>DAFTAR GAMBAR</b> .....                                | <b>xii</b>  |
| <b>DAFTAR TABEL</b> .....                                 | <b>xv</b>   |
| <b>BAB I PENDAHULUAN</b> .....                            | <b>1</b>    |
| A. Latar Belakang .....                                   | 1           |
| B. Rumusan Masalah.....                                   | 5           |
| C. Tujuan Penelitian.....                                 | 5           |
| D. Manfaat Penelitian .....                               | 6           |
| E. Batasan Masalah.....                                   | 7           |
| F. Penelitian Terkait.....                                | 8           |
| G. Kontribusi Utama Penelitian .....                      | 13          |
| <b>BAB II TINJAUAN PUSTAKA</b> .....                      | <b>14</b>   |
| A. Tersangka .....  | 14          |
| B. Wajah.....   | 14          |
| C. Visi Komputer.....                                     | 15          |
| D. Pengolahan Citra Digital.....                          | 16          |
| E. Jenis Citra Digital .....                              | 18          |
| F. Elemen Citra Digital .....                             | 20          |
| G. Resolusi Citra Digital .....                           | 22          |
| H. Video Digital .....                                    | 23          |
| I. Pengenalan Pola .....                                  | 24          |
| J. Deteksi Wajah.....                                     | 28          |
| K. Pengenalan Wajah .....                                 | 30          |
| L. Oklusi Pengenalan Wajah .....                          | 31          |
| M. Metode Viola-Jones .....                               | 31          |
| N. Support Vector Machine .....                           | 32          |
| O. Histogram of Oriented Gradient.....                    | 33          |
| P. Contrast Limited Adaptive Histogram Equalization ..... | 34          |
| Q. Convolutional Neural Network.....                      | 36          |
| R. Arsitektur CNN .....                                   | 41          |
| S. Cosine Similarity .....                                | 43          |
| T. Evaluasi Sistem .....                                  | 44          |
| U. Kerangka Pikir Penelitian .....                        | 45          |

|  |            |
|--|------------|
| <b>BAB III METODE PENELITIAN .....</b>   | <b>47</b>  |
| A. Rancangan Penelitian .....  | 47         |
| B. Jenis Penelitian.....   | 51         |
| C. Lokasi Dan Waktu Penelitian .....   | 51         |
| D. Sumber Data .....   | 52         |
| E. Instrumentasi Penelitian.....   | 54         |
| F. Perancangan Sistem .....  | 55         |
| <b>BAB IV HASIL DAN PEMBAHASAN.....</b>  | <b>77</b>  |
| A. Performa Pengenalan Wajah Dengan Jarak Yang Berbeda .....   | 77         |
| B. Performa Pengenalan Wajah Menggunakan Oklusi Kerudung .....   | 85         |
| C. Hasil Performa Pengenalan Wajah Untuk Resolusi Gambar Yang Berbeda-<br>Beda Termasuk Resolusi Rendah.....   | 94         |
| D. Perbandingan Performa <i>Histogram of Oriented Gradient - Multi-class<br/>    Support Vector Machine (HOG-MSVM)</i> dengan <i>Convolutional Neural<br/>    Network – Cosine Similarity (CNN-CS)</i> ..... | 163        |
| <b>BAB V KESIMPULAN DAN SARAN .....</b>  | <b>171</b> |
| A. Kesimpulan.....   | 171        |
| B. Saran.....  | 172        |
| <b>DAFTAR PUSTAKA.....</b>   | <b>173</b> |
| <b>BIODATA.....</b>  | <b>180</b> |
| <b>LAMPIRAN.....</b>   | <b>181</b> |

## DAFTAR GAMBAR

|              |  |    |
|--------------|--|----|
| Gambar 2.1.  | Koordinat Citra Digital .....  | 16 |
| Gambar 2.2.  | Representasi citra digital dalam dua dimensi .....   | 18 |
| Gambar 2.3.  | Citra biner dua warna .....  | 18 |
| Gambar 2.4.  | Citra biner empat warna.....   | 19 |
| Gambar 2.5.  | Citra biner delapan warna .....  | 19 |
| Gambar 2.6.  | Perbandingan gradasi warna mulai 2 bit - 10 bit .....  | 19 |
| Gambar 2.7.  | Penyimpanan citra warna di dalam memori .....  | 20 |
| Gambar 2.8.  | Algoritma pengenalan pola .....  | 25 |
| Gambar 2.9.  | Ilustrasi Deteksi Wajah .....  | 30 |
| Gambar 2.10. | Flowchart perhitungan metode viola-jones.....  | 32 |
| Gambar 2.11. | Ilustrasi distribusi excess piksel di bawah <i>cliplimit</i> .....   | 36 |
| Gambar 2.12. | Operasi Konvolusi.....   | 38 |
| Gambar 2.13. | Operasi max pooling .....  | 40 |
| Gambar 2.18. | Kerangka Pikir Penelitian .....  | 46 |
| Gambar 3.1.  | Tahapan Penelitian .....   | 48 |
| Gambar 3.2.  | Ilustrasi pengambilan data latih.....  | 52 |
| Gambar 3.3.  | Ilustrasi pengambilan data uji untuk skenario kasus pertama, kedua dan ketiga .....  | 53 |
| Gambar 3.4.  | Desain sistem untuk deteksi dan identifikasi wajah dengan jarak yang berbeda .....   | 56 |
| Gambar 3.5.  | Data latih 5 sudut.....  | 57 |
| Gambar 3.6.  | Contoh hasil deteksi fitur berdasarkan nilai <i>threshold</i> : (a) <i>threshold</i> = 0, (b) <i>threshold</i> = 1, (c). <i>threshold</i> = 5, (d) <i>threshold</i> = 10. .... | 58 |
| Gambar 3.7.  | Contoh <i>cropping</i> pada <i>pre-processing</i> .....  | 59 |
| Gambar 3.8.  | Resizing fitur (a) Wajah, (b) Sepasang Mata, (c) Hidung, (d) Mulut .....   | 60 |
| Gambar 3.9.  | Perbaikan citra (a) Citra masukan, (b) Citra luaran <i>brightness enhancement</i> , (c) Citra luaran <i>grayscaleing</i> .....   | 61 |
| Gambar 3.10. | Tahapan pada metode HOG .....  | 62 |
| Gambar 3.11. | Contoh pola fitur sepasang mata dengan ukuran 20 x 65 piksel.....  | 62 |
| Gambar 3.12. | Desain sistem identifikasi wajah dengan oklusi mengenakan kerudung .....   | 64 |
| Gambar 3.13. | Data latih tanpa mengenakan kerudung .....   | 66 |
| Gambar 3.14. | Data latih dengan mengenakan kerudung .....  | 67 |
| Gambar 3.15. | <i>Cropping</i> pada proses <i>Pre-processing</i> .....  | 68 |
| Gambar 3.16. | Peningkatan kualitas gambar (a) <i>input image</i> , (b) <i>output of brightness enhancement</i> , (c) <i>output of grayscaleing</i> .....                                     | 69 |
| Gambar 3.17. | Tahapan metode HOG pada kasus oklusi mengenakan kerudung.....  | 70 |
| Gambar 3.18. | Contoh fitur pola (a) Mulut, (b) Hidung, (c) Sepasang mata .....   | 71 |

|   |     |
|---|-----|
| Gambar 3.19. Desain sistem identifikasi wajah untuk resolusi yang berbeda-beda termasuk resolusi rendah.....  | 73  |
| Gambar 3.20. Gambar wajah data uji untuk data sumber .....  | 75  |
| Gambar 3.21. Gambar wajah data uji untuk data target.....   | 76  |
| Gambar 4.1. <i>Coding Matrix</i> pada model trainn .....  | 79  |
| Gambar 4.2. <i>Binary Learners</i> pada model trainn .....  | 80  |
| Gambar 4.3. Data latih tanpa mengenakan kerudung (TK) dan frame data uji tanpa mengenakan kerudung (TK) yang salah prediksi pada target kedua (T2) .....  | 92  |
| Gambar 4.4. Performa pengenalan wajah untuk gambar yang berbeda-beda termasuk resolusi rendah.....  | 94  |
| Gambar 4.5. Data sumber dan tiga data target yang gagal dikenali pada gambar dengan ukuran 512 piksel .....   | 97  |
| Gambar 4.6. Data Sumber dan empat Data Target yang gagal dikenali pada gambar dengan ukuran 256 piksel .....  | 111 |
| Gambar 4.7. Data Sumber dan empat Data Target yang gagal dikenali pada gambar dengan ukuran 128 piksel .....  | 124 |
| Gambar 4.8. Data Sumber dan 11 Data Target yang gagal dikenali pada gambar dengan ukuran 64 piksel .....  | 138 |
| Gambar 4.9. Data Sumber dan 20 data target yang gagal dikenali pada gambar dengan ukuran 32 piksel .....  | 152 |
| Gambar 4.10. Performa pengenalan wajah untuk HOG-MSVM untuk kondisi jarak yang berbeda. ....  | 164 |
| Gambar 4.11. Performa pengenalan wajah untuk HOG-MSVM untuk kondisi jarak yang berbeda. ....  | 164 |
| Gambar 4.12. Performa HOG-MSVM pada data latih tidak mengenakan kerudung dan uji tidak mengenakan kerudung dibandingkan dengan data latih tidak mengenakan kerudung dan data uji dengan mengenakan kerudung. ....   | 165 |
| Gambar 4.13. Performa HOG-MSVM pada data latih dengan mengenakan kerudung dan uji tidak mengenakan kerudung dibandingkan dengan data latih dengan mengenakan kerudung dan data uji dengan mengenakan kerudung. .... | 166 |
| Gambar 4.14. Performa HOG-MSVM pada data latih tidak mengenakan kerudung dan uji tidak mengenakan kerudung dibandingkan dengan data latih tidak mengenakan kerudung dan data uji dengan mengenakan kerudung. ....   | 167 |
| Gambar 4.15. Performa HOG-MSVM pada data latih dengan mengenakan kerudung dan uji tidak mengenakan kerudung dibandingkan dengan data latih dengan mengenakan kerudung dan data uji dengan mengenakan kerudung. .... | 167 |

|  |     |
|--|-----|
| Gambar 4.16. Perbandingan Performa HOG-MSVM VS CNN-CS<br>untuk kondisi data latih TK dan data uji TK ..... | 169 |
| Gambar 4.17. Perbandingan Performa HOG-MSVM VS CNN-CS<br>untuk kondisi data latih TK dan data uji DK.....  | 169 |
| Gambar 4.18. Perbandingan Performa HOG-MSVM VS CNN-CS<br>untuk kondisi data latih DK dan data uji TK.....  | 170 |
| Gambar 4.19. Perbandingan Performa HOG-MSVM VS CNN-CS<br>untuk kondisi data latih DK dan data uji DK ..... | 170 |

## DAFTAR TABEL

|             |   |     |
|-------------|---|-----|
| Tabel 1.1.  | <i>State of the Art</i> .....   | 9   |
| Tabel 4.1.  | Hasil Pengujian jarak 300 cm untuk pengenalan wajah pada target kedua (T2).....   | 81  |
| Tabel 4.2.  | <i>Confusion matrix</i> pengenalan wajah untuk jarak 200 cm .....   | 82  |
| Tabel 4.3.  | <i>Confusion matrix</i> pengenalan wajah untuk jarak 300 cm .....   | 83  |
| Tabel 4.4.  | <i>Confusion matrix</i> pengenalan wajah untuk jarak 400 cm .....   | 84  |
| Tabel 4.5.  | Hasil pengujian skenario pertama (TK-TK) dan skenario kedua (TK-DK).....  | 86  |
| Tabel 4.6.  | Hasil pengujian skenario ketiga (DK-TK) dan skenario keempat (DK-DK).....   | 86  |
| Tabel 4.7.  | <i>Confusion matrix</i> pengenalan wajah untuk data latih tidak mengenakan kerudung (TK) dan data uji tidak mengenakan kerudung (TK). .....                         | 88  |
| Tabel 4.8.  | <i>Confusion matrix</i> pengenalan wajah untuk data latih tidak mengenakan kerudung (TK) dan data uji dengan mengenakan kerudung (DK). .....                        | 88  |
| Tabel 4.9.  | <i>Confusion matrix</i> pengenalan wajah untuk data latih dengan mengenakan kerudung (DK) dan data uji tidak mengenakan kerudung (TK). .....                        | 89  |
| Tabel 4.10. | <i>Confusion matrix</i> pengenalan wajah untuk data latih dengan mengenakan kerudung (DK) dan data uji dengan mengenakan kerudung (DK). .....                       | 89  |
| Tabel 4.11. | Hasil pengujian untuk data latih tanpa mengenakan kerudung (TK) dan data uji tanpa mengenakan kerudung (TK) untuk target kedua (T2).....                            | 90  |
| Tabel 4.12. | Hasil pengujian untuk data latih tanpa mengenakan kerudung (TK) dan data uji tanpa mengenakan kerudung (TK) untuk target kedua (T2) dengan status salah prediksi .. | 91  |
| Tabel 4.13. | Hasil performa pengenalan wajah untuk gambar dengan resolusi 512 piksel .....   | 95  |
| Tabel 4.14. | Hasil Pengujian dengan menggunakan ukuran 512 piksel untuk target dengan ID000.....   | 95  |
| Tabel 4.15. | Hasil Pengujian dengan menggunakan ukuran 512 piksel untuk target dengan ID001.....   | 98  |
| Tabel 4.16. | Hasil Pengujian dengan menggunakan ukuran 512 piksel untuk target dengan ID002.....   | 99  |
| Tabel 4.18. | Hasil Pengujian dengan menggunakan ukuran 512 piksel untuk target dengan ID007.....   | 101 |
| Tabel 4.20. | Hasil Pengujian dengan menggunakan ukuran 512 piksel untuk target dengan ID011.....   | 103 |
| Tabel 4.21. | Hasil Pengujian dengan menggunakan ukuran 512 piksel untuk target dengan ID012.....   | 105 |
| Tabel 4.22. | Hasil Pengujian dengan menggunakan ukuran 512 piksel untuk target dengan ID015.....   | 106 |

|  |     |
|--|-----|
| Tabel 4.23. Hasil Pengujian dengan menggunakan ukuran 512 piksel untuk target dengan ID018.....  | 107 |
| Tabel 4.24. Hasil performa pengenalan wajah untuk gambar dengan resolusi <b>256</b> piksel ..... | 109 |
| Tabel 4.25. Hasil Pengujian dengan menggunakan ukuran 256 piksel untuk target dengan ID000.....  | 109 |
| Tabel 4.26. Hasil Pengujian dengan menggunakan ukuran 256 piksel untuk target dengan ID001.....  | 111 |
| Tabel 4.27. Hasil Pengujian dengan menggunakan ukuran 256 piksel untuk target dengan ID002.....  | 112 |
| Tabel 4.28. Hasil Pengujian dengan menggunakan ukuran 256 piksel untuk target dengan ID006.....  | 113 |
| Tabel 4.29. Hasil Pengujian dengan menggunakan ukuran 256 piksel untuk target dengan ID007.....  | 115 |
| Tabel 4.30. Hasil Pengujian dengan menggunakan ukuran 256 piksel untuk target dengan ID008.....  | 116 |
| Tabel 4.31. Hasil Pengujian dengan menggunakan ukuran 256 piksel untuk target dengan ID011.....  | 117 |
| Tabel 4.32. Hasil Pengujian dengan menggunakan ukuran 256 piksel untuk target dengan ID012.....  | 118 |
| Tabel 4.33. Hasil Pengujian dengan menggunakan ukuran 256 piksel untuk target dengan ID015.....  | 119 |
| Tabel 4.34. Hasil Pengujian dengan menggunakan ukuran 256 piksel untuk target dengan ID018.....  | 120 |
| Tabel 4.35. Hasil performa pengenalan wajah untuk gambar dengan resolusi <b>128</b> piksel ..... | 122 |
| Tabel 4.36. Hasil Pengujian dengan menggunakan ukuran 128 piksel untuk target dengan ID000.....  | 122 |
| Tabel 4.37. Hasil Pengujian dengan menggunakan ukuran 128 piksel untuk target dengan ID001.....  | 124 |
| Tabel 4.38. Hasil Pengujian dengan menggunakan ukuran 128 piksel untuk target dengan ID002.....  | 126 |
| Tabel 4.39. Hasil Pengujian dengan menggunakan ukuran 128 piksel untuk target dengan ID006.....  | 127 |
| Tabel 4.40. Hasil Pengujian dengan menggunakan ukuran 128 piksel untuk target dengan ID007.....  | 128 |
| Tabel 4.41. Hasil Pengujian dengan menggunakan ukuran 128 piksel untuk target dengan ID008.....  | 129 |
| Tabel 4.42. Hasil Pengujian dengan menggunakan ukuran 128 piksel untuk target dengan ID011.....  | 130 |
| Tabel 4.43. Hasil Pengujian dengan menggunakan ukuran 128 piksel untuk target dengan ID012.....  | 131 |
| Tabel 4.44. Hasil Pengujian dengan menggunakan ukuran 128 piksel untuk target dengan ID015.....  | 132 |
| Tabel 4.45. Hasil Pengujian dengan menggunakan ukuran 128 piksel untuk target dengan ID018.....  | 134 |

|   |     |
|---|-----|
| Tabel 4.46. Hasil performa pengenalan wajah untuk gambar dengan resolusi 64 piksel .....              | 135 |
| Tabel 4.47. Hasil Pengujian dengan menggunakan ukuran <b>64</b> piksel untuk target dengan ID000..... | 136 |
| Tabel 4.48. Hasil Pengujian dengan menggunakan ukuran 64 piksel untuk target dengan ID001.....        | 138 |
| Tabel 4.49. Hasil Pengujian dengan menggunakan ukuran 64 piksel untuk target dengan ID002.....        | 139 |
| Tabel 4.50. Hasil Pengujian dengan menggunakan ukuran 64 piksel untuk target dengan ID006.....        | 140 |
| Tabel 4.51. Hasil Pengujian dengan menggunakan ukuran 64 piksel untuk target dengan ID007.....        | 142 |
| Tabel 4.52. Hasil Pengujian dengan menggunakan ukuran 64 piksel untuk target dengan ID008.....        | 143 |
| Tabel 4.53. Hasil Pengujian dengan menggunakan ukuran 64 piksel untuk target dengan ID011.....        | 144 |
| Tabel 4.54. Hasil Pengujian dengan menggunakan ukuran 64 piksel untuk target dengan ID012.....        | 145 |
| Tabel 4.55. Hasil Pengujian dengan menggunakan ukuran 64 piksel untuk target dengan ID015.....        | 146 |
| Tabel 4.56. Hasil Pengujian dengan menggunakan ukuran 64 piksel untuk target dengan ID018.....        | 148 |
| Tabel 4.57. Hasil performa pengenalan wajah untuk gambar dengan resolusi 32 piksel .....              | 149 |
| Tabel 4.58. Hasil Pengujian dengan menggunakan ukuran <b>32</b> piksel untuk target dengan ID000..... | 150 |
| Tabel 4.59. Hasil Pengujian dengan menggunakan ukuran 32 piksel untuk target dengan ID001.....        | 152 |
| Tabel 4.60. Hasil Pengujian dengan menggunakan ukuran 32 piksel untuk target dengan ID002.....        | 154 |
| Tabel 4.61. Hasil Pengujian dengan menggunakan ukuran 32 piksel untuk target dengan ID006.....        | 155 |
| Tabel 4.62. Hasil Pengujian dengan menggunakan ukuran 32 piksel untuk target dengan ID007.....        | 156 |
| Tabel 4.63. Hasil Pengujian dengan menggunakan ukuran 32 piksel untuk target dengan ID008.....        | 157 |
| Tabel 4.64. Hasil Pengujian dengan menggunakan ukuran 32 piksel untuk target dengan ID011.....        | 158 |
| Tabel 4.65. Hasil Pengujian dengan menggunakan ukuran 32 piksel untuk target dengan ID012.....        | 159 |
| Tabel 4.66. Hasil Pengujian dengan menggunakan ukuran 32 piksel untuk target dengan ID015.....        | 161 |
| Tabel 4.67. Hasil Pengujian dengan menggunakan ukuran 32 piksel untuk target dengan ID018.....        | 162 |

# **BAB I**

## **PENDAHULUAN**

### **A. Latar Belakang**

Indonesia merupakan negara kepulauan terbesar di dunia, dengan jumlah sekitar 17.508 pulau yang berada diantara benua Asia dan Australia. Banyaknya pulau dan letaknya yang memisahkan dua benua, diperlukan transportasi udara untuk membangun konektivitas dan memobilisasi daerah-daerah tersebut guna pemerataan pembangunan dan pemeliharaan keutuhan Negara Kesatuan Republik Indonesia (NKRI) (Gunawan & Medianto, 2017).

Berdasarkan data dari Direktorat Jenderal Perhubungan Udara, jumlah bandar udara (bandara) di Indonesia menurut penggunaan bandara pada tahun 2021 yakni 320 untuk domestik dan 31 untuk Internasional. Keberadaan bandara sebagai sarana pendukung transportasi udara yang berperan sebagai penunjang, pendorong dan penggerak bagi pembangunan suatu daerah. Salah satu fasilitas keamanan penerbangan yang harus terpasang pada bandara khususnya pada area pemeriksaan keberangkatan penumpang adalah *Closed Circuit Television* (CCTV). CCTV sangat penting keberadaannya untuk memantau pergerakan atau segala aktivitas manusia, namun penggunaannya saat ini hanya sebatas memonitor hasil tangkapan CCTV pada layar dan melihat hasil rekaman jika diperlukan. Padahal keberadaan CCTV tersebut bisa dioptimalkan untuk mendeteksi tersangka atau daftar pencarian orang (DPO), baik untuk yang

terduga sebagai pelaku aksi teror atau pelaku kriminal (Badan Pusat Statistik, 2016), (Perhubungan, 2018).

Mabes Polri merilis laporan akhir tahun, pada tanggal 22 Agustus 2020. Kapolri Jenderal Idham Azis menginformasikan bahwa terdapat 228 tersangka kasus terorisme yang ditangkap oleh Polri, selain itu terdapat penangkapan 23 teroris, dimana di dalamnya terdapat 2 orang yang diduga terlibat dalam aksi teror di Poso, Sulawesi Tengah hingga kasus bom Bali 1. Idham Azis juga membahas tentang penegakan hukum terhadap 2 DPO kelompok Ali Kalora di Poso, yakni Aziz Arifin dan Wahid, keduanya ditembak mati dalam pengejaran polisi pada pertengahan November 2020 (Adyatama, 2020).

Untuk mengantisipasi terjadinya tindakan terorisme, maka pemerintah diamanahkan oleh Undang Undang (UU) No 5 tahun 2018 tentang Pemberantasan Tindak Pidana Terorisme pada pasal 43A yang berbunyi: Pemerintah wajib melakukan pencegahan tindak pidana terorisme; dalam upaya pencegahan tindak pidana terorisme, pemerintah melakukan langkah antisipasi secara terus menerus yang dilandasi dengan prinsip perlindungan hak asasi manusia dan prinsip kehati-hatian (Indonesia, 2018). Selain itu, pada Rencana Induk Riset Nasional (RIRN) 2017-2045, ditegaskan bahwa pembangunan kekuatan pertahanan negara Indonesia merupakan suatu kebutuhan yang tidak dapat dihindari. Mencermati dinamika konteks strategis, ancaman yang sangat mungkin dihadapi Indonesia ke depan dapat berbentuk ancaman konvensional dan ancaman

non-konvensional, baik yang bersumber dari luar negeri maupun dari dalam negeri. Beberapa ancaman yang perlu diwaspadai saat ini adalah terorisme, gerakan separatisme dan aksi radikalisme (Riset & Tinggi, 2017).

Pelaku tindak kriminal berdasarkan pada data registrasi Polri, mencatat bahwa selama periode tahun 2017–2019, jumlah kejahatan pada tahun 2017 sebanyak 336.652 kejadian, lalu tahun 2018 sebanyak 294.281 kejadian dan pada tahun 2019 sekitar 269.324 kejadian. Jenis kejadian pencurian merupakan kejahatan yang paling banyak terjadi, jumlahnya mencapai lebih dari 36-45 persen (Keamanan, 2020).

Menurut Hierarki Kebutuhan Maslow (*Maslow's Hierarchy of Needs*), kebutuhan manusia dapat dikategorikan menjadi 5 level: kebutuhan fisiologi, kebutuhan akan keamanan, kebutuhan akan memiliki dan kasih sayang, kebutuhan akan penghargaan, dan kebutuhan akan aktualisasi diri. Berdasarkan hasil riset Abraham Maslow, kebutuhan ini sifatnya bertingkat, sebelum bisa memuaskan kebutuhan di level berikutnya, maka kebutuhan di level sebelumnya harus dipenuhi dulu. Setelah memenuhi kebutuhan level pertama, selanjutnya bisa mulai mewujudkan kebutuhan level kedua yaitu *safety needs* atau kebutuhan akan keamanan. Konten dari kebutuhan keamanan adalah keamanan fisik dan keamanan nonfisik diantaranya yakni stabilitas dan ketergantungan, perlindungan dari kriminalitas, terorisme, perang, penyakit, bencana alam, dan lain – lain. Keamanan juga termasuk keamanan secara mental seperti kebebasan dari rasa takut, stres, dll (Iskandar, 2019).

Mengacu pada fakta tersebut maka aspek keamanan sebagai level kedua sangat dibutuhkan oleh manusia, sementara bandara sebagai gerbang yang banyak dilewati untuk meninggalkan atau mendatangi suatu daerah termasuk orang yang terkategori sebagai tersangka atau DPO. Oleh karena itu, penggunaan kamera CCTV yang memiliki fungsi untuk melakukan monitoring dioptimalkan dengan memanfaatkan teknologi pengenalan wajah yang saat ini mendapat perhatian sangat besar pada bidang keamanan yang berbasis sistem biometrika. Sistem biometrika merupakan teknologi untuk pengenalan diri dengan menggunakan bagian tubuh atau perilaku manusia (Son et al., 2020).

Pengenalan wajah berbasis video sebagai salah satu bagian dari sistem biometrika telah menjadi perhatian luas dan merupakan salah satu topik penting pada penelitian di bidang *image processing* untuk mengidentifikasi wajah seseorang (Raghuwanshi, 2017). Pengenalan wajah dengan gambar wajah masih banyak dipelajari, sementara penelitian pada pengenalan wajah berbasis video relatif tidak memadai (Huang et al., 2015). Wajah berbasis video, ekspresi dan pengenalan adegan adalah masalah mendasar dalam interaksi manusia-mesin, terutama ketika video berdurasi pendek (Hajati et al., 2017). Tampilan wajah yang ditangkap dengan kamera video bervariasi secara signifikan karena perubahan dalam pose, iluminasi, skala, blur, ekspresi dan oklusi (Parchami et al., 2017). Selain itu wajah manusia dalam video pengawasan sering mengalami gambar buram, variasi pose dramatis dan oklusi (Ding & Tao, 2018).

Penelitian yang telah dilakukan ini memanfaatkan kamera CCTV sebagai media untuk mengambil data uji dan kamera DSLR (*Digital Single-Lens Reflex*) untuk mengambil data latih, kemudian menggunakan kecerdasan buatan dalam pembuatan sistem untuk melakukan deteksi dan identifikasi wajah seseorang guna mengantisipasi secara dini kehadiran tersangka pelaku teror, gerakan separatisme, aksi radikalisme dan tindakan kriminal. Inilah yang menjadi alasan kehadiran penelitian dengan judul “Deteksi Tersangka Dengan Pengenalan Wajah Berbasis Visi Komputer (*Face Recognition for Suspect Detection Using Computer Vision*)” sebagai langkah antisipasi mendeteksi tersangka atau DPO.

### **B. Rumusan Masalah**

Berdasarkan pada uraian latar belakang, maka rumusan masalah pada penelitian ini yakni:

1. Bagaimana mengenali tersangka menggunakan visi komputer dengan kondisi jarak yang berbeda, oklusi dengan mengenakan kerudung dan kondisi resolusi gambar yang berbeda-beda termasuk resolusi rendah?
2. Bagaimana kinerja sistem dalam mengenali tersangka menggunakan visi komputer dengan kondisi jarak yang berbeda, oklusi dengan mengenakan kerudung dan kondisi resolusi gambar yang berbeda-beda termasuk resolusi rendah?

### **C. Tujuan Penelitian**

Tujuan yang telah dicapai pada penelitian ini dengan mengacu pada rumusan masalah yakni:

1. Merancang sistem untuk mengenali tersangka menggunakan visi komputer dengan kondisi jarak yang berbeda, oklusi dengan mengenakan kerudung dan kondisi resolusi gambar yang berbeda-beda termasuk resolusi rendah.
2. Menampilkan hasil kinerja sistem dalam mengenali tersangka menggunakan visi komputer dengan kondisi jarak yang berbeda, oklusi dengan mengenakan kerudung dan kondisi resolusi gambar yang berbeda-beda termasuk resolusi rendah.

#### **D. Manfaat Penelitian**

Berdasarkan pada rumusan masalah dan tujuan penelitian, penulis mengharapkan agar hasil dari penelitian ini memberikan manfaat minimal sebagai berikut:

1. Menambah khazanah ilmu pengetahuan pada bidang penelitian tentang deteksi dan pengenalan wajah yang berbasis pada kecerdasan buatan.
2. Berkontribusi pada penyediaan sistem yang dapat melakukan deteksi dan pengenalan wajah, dimana pada penelitian ini titik fokusnya mengarah pada deteksi tersangka pelaku teror atau kriminal yang masuk daftar pencarian orang (DPO).
3. Memberikan kemudahan dalam melakukan monitoring secara berkelanjutan dengan melakukan deteksi dan pengenalan wajah pelaku teror atau kriminal yang masuk DPO dan terdapat pada antrian, tanpa mengganggu penumpang yang lainnya pada area pemeriksaan keberangkatan penumpang di bandara.

### **E. Batasan Masalah**

Pada penelitian ini, terdapat beberapa batasan masalah diantaranya adalah sebagai berikut:

1. Data latih yang digunakan berupa foto, sementara untuk data ujinya bersumber dari data video yang menggunakan kamera CCTV.
2. Pengambilan data uji diskenariokan seperti kondisi area pemeriksaan penumpang pesawat terbang pada terminal keberangkatan.
3. Terdapat 3 kasus pada penelitian ini yakni:
  - a) Deteksi dan identifikasi wajah dengan jarak berbeda yakni, 200 cm, 300 cm dan 400 cm.
  - b) Deteksi dan identifikasi wajah menggunakan oklusi dengan mengenakan kerudung.
  - c) Deteksi dan identifikasi wajah dengan ukuran resolusi gambar yang berbeda yakni, 512, 256, 128, 64 dan 32 piksel.
4. Perbandingan data latih dan data uji yang digunakan pada kasus pertama adalah 75 data latih dan 2.250 data uji, pada kasus kedua total data latih yang digunakan adalah 75, sementara total data uji yang digunakan adalah 1.500. Selanjutnya pada kasus ketiga total data latih yang digunakan adalah 490.623, sedangkan total data uji yang digunakan adalah 2.000.

## F. Penelitian Terkait

Pada bagian ini dibahas mengenai penelitian yang terkait dengan deteksi dan identifikasi wajah:

1. Penelitian yang dilakukan oleh Manjeet Singh dan Ravi Sahran pada tahun 2018 dengan judul "*A New Approach for Suspect Detection in Video Surveillance*", menggunakan metode Viola Jones untuk deteksi wajah dan metode *eigenface* untuk pengenalan wajah. Adapun nilai akurasi yang didapatkan yakni 75,69% untuk deteksi wajah dan 70,65% untuk identifikasi wajah (Singh & Sahran, 2018).
2. Penelitian dengan judul "*The excellent properties of a dense grid-based HOG feature on face recognition compared to Gabor and LBP*" yang dilakukan pada tahun 2018 oleh Zheng Xiang dkk, pada penelitian ini membandingkan 3 metode dalam pengenalan wajah yakni HOG, Gabor dan LBP. Hasil penelitian menunjukkan bahwa metode HOG memiliki akurasi yang lebih baik jika dibandingkan dengan LBP dan Gabor, dimana nilai akurasi pengenalan wajah yang didapatkan pada HOG sebesar 61,1%, LBP sebesar 57,7 % dan Gabor sebesar 49,1% (Xiang et al., 2018).
3. Ade Nurhopipah dan Agus Harjoko, pada tahun 2018 melakukan penelitian dengan judul "*Motion Detection and Face Recognition for CCTV Surveillance System*", pada penelitian ini terdapat tiga keluaran yakni informasi deteksi gerak, kemudian informasi deteksi wajah dan informasi identifikasi wajah. Hasil akurasi dari deteksi

wajah sebesar 76% dengan menggunakan metode *Haar Cascade Classifier* sedangkan hasil identifikasi wajah sebesar 60% dengan menggunakan *Speeded-UP Robust Features (SURF)* dan *Principal Component Analysis (PCA)* pada proses ekstraksi fiturnya (Nurhopipah & Harjoko, 2018).

4. Pada penelitian yang dilakukan oleh V.D. Ambeth Kumar dkk yang berjudul "*Facial Recognition System for Suspect Identification Using a Surveillance Camera*", tahun 2018. Penelitian ini menggunakan metode Viola Jones untuk deteksi wajah, sementara untuk identifikasi wajah menggunakan metode *Local Binary Pattern Histogram (LBPH)*. Untuk data latih menggunakan foto dengan empat angle yang berbeda dengan 3 orang sebagai target. Hasil penelitian menunjukkan bahwa dari 3 orang target yang digunakan, berhasil dikenali 2 target dan 1 target tidak dikenali (Kumar et al., 2018).

Selain literatur tersebut, terdapat tambahan penelitian terdahulu yang menjadi rujukan pada penelitian ini, sebagaimana yang ditampilkan pada tabel 1.1. berikut ini.

Tabel 1.1. *State of the Art*

| No | Peneliti/<br>Tahun  | Judul/Tahun   | Media<br>Pengenalan                       | Metode  | Hasil                             |
|----|---|---|---|---|-----------------------------------|
| 1  | Yuan Gao,<br>Jiayi Ma,<br>and Alan<br>L. Yuille,<br>Fellow,<br>IEEE<br>(2017) | <i>Semi-Supervised Sparse Representation Based Classification for Face Recognition with</i> | Data latih dan data uji menggunakan photo | Viola Jones<br>Deep Convolution Neural Network (DCNN)<br>Semi-Supervised Sparse | Akurasi yang dicapai yaitu 97,74% |

| No | Peneliti/<br>Tahun   | Judul/Tahun   | Media<br>Pengenalan   | Metode   | Hasil   |
|----|--|---|---|--|---|
|    |  | <i>Insufficient Labeled Samples</i> (Gao et al., 2017)  |   | Representation (S <sup>3</sup> RC) dan Variation Dictionary Learned (SVDL)                                     |   |
| 2  | Xiang Chen, Limbo Qing, Xiaohai he, Jie Su, Yonghong Peng (2018) | <i>From Eyes to Face Synthesis: A New Approach for Human – Centered Smart Surveillance</i> (Chen et al., 2018)                          | Data latih menggunakan Photo dan data uji menggunakan video                 | Generative Adversarial Networks (GAN)  | Pada penelitian ini diperoleh akurasi lebih dari 91%  |
| 3  | Mohannad A. Abuzneid, Ausif Mahmood (2018)                       | <i>Enhanced human face recognition using LBPH descriptor, multi-kkn, and back propagation neural network</i> (Abuzneid & Mahmood, 2018) | Untuk data latih dan data ujinya menggunakan data photo                     | Local binary pattern (LBPH), Multi-KKN dan Back Propagation neural network.                                    | Akurasi yang didapatkan pada penelitian ini sebesar 95.71%  |
| 4  | Alexey D. Egorov, dkk (2018)                                     | <i>Some Cases of Optimization Face Detection Methodes on Image, Using the Viola-Jones Method as an Example</i> (Egorov et al., 2018)    | Data latih dan data uji dari kamera digital (foto)                          | Deteksi Wajah dengan Viola Jones (membandingkan komputasi waktu berdasarkan cascade yang ada pada viola jones) | Akurasinya dapat mengurangi komputasi waktu hingga 7,5938 kali.   |
| 5  | Wen-Hui Lin, Ping Wang, Chen-Fang Tsai (2016)                    | <i>Face recognition using support vector model classifier for user authentication</i> (Lin et al., 2016)                                | Data latih dan uji foto   | machine learning and the generalization capability of support vector models (SVMs)                             | Pengenalan wajah lebih dari 97% dengan skema validasi silang dengan resolusi 168x341 dan sekitar 89% secara rata2 dengan resolusi beragam |
| 6  | Edwin Jose, Greeshma M., Mithun Haridas (2019)                   | <i>Face Recognition based Surveillance System Using</i>   | data latih yang digunakan berupa foto, sementara untuk data uji menggunakan | FaceNet dan MTCNN pada Jetson TX2  | Deteksi akurasi mencapai 97%  |

| No | Peneliti/<br>Tahun  | Judul/Tahun  | Media<br>Pengenalan   | Metode  | Hasil  |
|----|---|--|---|---|--|
|    |   | <i>FaceNet and MTCNN on Jetson TX2</i> (Jose et al., 2019)   | data foto dan video   |   |  |
| 7  | Divyarajsinh N. Parmar, Brijesh B. Mehta (2013)                                 | <i>Face Recognition Methods &amp; Applications</i> (Parmar & Mehta, 2014)  | Data latih dan uji Foto   | holistic matching method, feature extraction method and hybrid methods  | Implementasi face recognition pada sebuah aplikasi   |
| 8  | Abhishek Jha, monika hooda (2013)   | <i>ClassRoom Attendance System Using Facial Recognition System</i> (Abhishek Jha, 2007)                                    | Data latih dan data uji menggunakan photo   | face detection, Face Alignment, Feature Extraction, Face Matching   | Implementasi face recognition pada absensi   |
| 9  | Lacey Best-Rowden, Hu Han, Charles Otto, Brendan Klare, and Anil K. Jain (2014) | <i>Unconstrained Face Recognition: Identifying a Person of Interest from a Media Collection</i> (Best-Rowden et al., 2014) | data latih foto dan data uji foto dan video   | commercial off the shelf (COTS)   | Peningkatan akurasi dengan commercial off the shelf (COTS) pada kualitas rendah  |
| 10 | Muhammad Arafah   | Deteksi Tersangka dengan Pengenalan Wajah Berbasis Visi Komputer   | Pada penelitian ini untuk data latihnya menggunakan data photo, sedangkan untuk data ujinya menggunakan data video. | Metode yang digunakan untuk deteksi wajah yakni Viola Jones, untuk ekstraksi fitur menggunakan Histogram of Oriented Gradients (HOG), Convolutional Neural Network, sementara untuk klasifikasinya menggunakan Multi-class Support Vector Machine (MSVM) dan Cosine similarity. | Penelitian ini dilakukan dengan tiga studi kasus dengan hasil:<br>1. Pengenalan wajah dengan jarak yang berbeda yakni; 200, 300 dan 400 cm, menggunakan metode Viola Jones, Brightness Enhancement, grayscling, Histogram Equalization, HOG dan MSVM nilai akurasi terbaik pada jarak 300 cm yakni 86,77%.<br>2. Pengenalan wajah dengan oklusi berupa kerudung dengan metode yang sama pada kasus pertama, menggunakan 4 skenario dengan nilai akurasi 85% pada kondisi data latih TK dan data uji TK, 61,40% pada kondisi data latih TK dan data |

| No | Peneliti/<br>Tahun | Judul/Tahun | Media<br>Pengenalan | Metode | Hasil   |
|----|--------------------|-------------|---------------------|--------|---|
|    |                    |             |                     |        | uji DK, 58,60% pada kondisi data latih DK dan data uji TK, dan 81,40% pada kondisi data latih DK dan data uji DK.<br>3. Pengenalan wajah untuk gambar dengan ukuran resolusi gambar yang berbeda yakni 512, 256, 128, 64 dan 32 piksel, menggunakan metode CNN dan Cosine Similarity. Pada resolusi gambar yang terkategori rendah, sistem dapat mengenali dengan baik, yakni menghasilkan akurasi sebesar 99,35% untuk gambar dengan resolusi 64 piksel dan 97,90% untuk gambar dengan resolusi 32 piksel. |

Optimalisasi pemanfaatan CCTV guna melakukan monitoring secara berkelanjutan yang diintegrasikan dengan adanya sistem pengenalan wajah di dalamnya untuk mengidentifikasi setiap wajah orang yang melewati area pemeriksaan penumpang pada terminal keberangkatan. Hal ini merupakan salah satu solusi yang bisa dilakukan saat ini, sebagai antisipasi keterbatasan manusia dalam melakukan pemantauan dan monitoring terhadap setiap orang yang melewati area pemeriksaan penumpang, baik secara langsung maupun melalui layar keluaran hasil tangkapan dari CCTV.

*State of the art* dari penelitian ini adalah tersedianya sistem untuk mengenali tersangka menggunakan visi komputer dengan jarak terbaik

untuk kasus identifikasi wajah pada area pemeriksaan penumpang di terminal keberangkatan, kemudian sistem juga dapat melakukan identifikasi wajah dengan kondisi terdapat oklusi mengenakan kerudung. Selain itu pada penelitian ini dilakukan juga pengujian performa sistem pengenalan wajah pada kondisi CCTV yang memiliki resolusi gambar yang berbeda-beda, termasuk kondisi gambar dengan resolusi rendah.

### **G. Kontribusi Utama Penelitian**

Pada bagian ini, akan dikemukakan beberapa kontribusi utama dari penelitian, diantaranya adalah:

1. Rekomendasi penentuan jarak terbaik untuk pemasangan CCTV pada area pemeriksaan penumpang di bandara. Penentuan jarak ini digunakan untuk kebutuhan deteksi dan pengenalan wajah tersangka berbasis visi komputer.
2. Peningkatan akurasi pengenalan wajah dengan menggunakan *pre-processing* berupa *Cropping*, *Resize*, *Brightness Enhancement*, *Grayscale* dan *Histogram Equalization* berbasis visi komputer pada jarak yang berbeda dan kondisi oklusi dimana tersangka mengenakan kerudung.
3. Pengujian model *embedding* berbasis CNN dan CS pada citra wajah yang terdeteksi dari kamera CCTV. Pengujian dilakukan pada Citra wajah dengan ukuran 512, 256, 128, 64 dan 32 piksel untuk melihat kekurangan dan kelebihan model *embedding* dalam mengklasifikasikan citra wajah dengan ukuran berbeda-beda.

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **A. Tersangka**

Tersangka berdasarkan Kamus Besar Bahasa Indonesia (KBBI) dapat diartikan sebagai diduga atau dicurigai, selain itu dapat juga didefinisikan sebagai orang yang telah disangka berdasarkan keterangan saksi atau pengakuannya sendiri (Badan Pengembangan dan Pembinaan Bahasa, 2016a). Menurut UU No. 8 Tahun 1981, pada pasal ke-1 poin ke-14, tersangka adalah seorang yang karena perbuatannya atau keadaannya, berdasarkan bukti permulaan patut diduga sebagai pelaku tindak pidana (Undang-Undang (UU) Tentang Hukum Acara Pidana, 1981).

Tindak pidana menurut S. R. Sianturi merupakan suatu tindakan pada tempat, waktu dan keadaan tertentu, yang dilarang (atau melanggar keharusan) dan diancam dengan pidana oleh undang-undang serta bersifat melawan hukum dan mengandung unsur kesalahan yang dilakukan oleh seseorang yang mampu bertanggung jawab (hukumonline, 2020). Salah satu tindak pidana yang banyak diberitakan beberapa tahun terakhir ini adalah tindak pidana terorisme, dan secara spesifik dibuatkan aturan secara rinci, sebagaimana yang termaktub pada Undang-Undang Republik Indonesia Nomor 5 tahun 2018 tentang Pemberantasan Tindak Pidana Terorisme.

#### **B. Wajah**

Berdasarkan KBBI (Kamus Besar Bahasa Indonesia) secara etimologi wajah adalah bagian depan dari kepala; roman muka; muka.

Wajah pada manusia menjadi fokus perhatian di dalam interaksi sosial, selain itu wajah memiliki peran penting dalam mengidentifikasi seseorang. Adapun batasan wajah pada manusia meliputi wilayah antara dahi dan dagu termasuk rambut, alis, pelipis, mata, hidung, pipi, mulut, bibir, gigi, kulit, dahi dan dagu (Badan Pengembangan dan Pembinaan Bahasa, 2016b).

### **C. Visi Komputer**

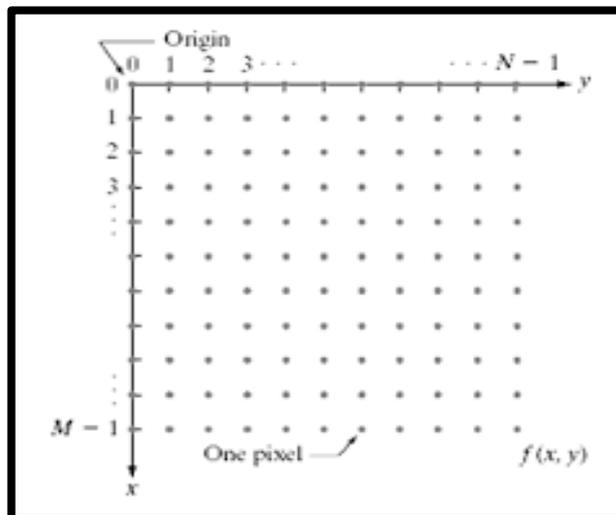
Visi Komputer merupakan suatu pembelajaran menganalisis gambar dan video untuk memperoleh hasil sebagaimana yang bisa dilakukan manusia. Sejatinya, visi komputer mencoba meniru cara kerja sistem visual manusia. Visi komputer dapat diibaratkan dengan penglihatan manusia yang dapat melihat objek dengan indera penglihatan (mata), selanjutnya citra objek diteruskan ke fungsi otak untuk diinterpretasi sehingga manusia mengerti objek apa yang tampak dalam penglihatan matanya (Masithoh et al., 2012).

Sistem visi komputer memproses gambar yang diperoleh dari kamera elektronik, dimana ilustrasinya seperti sistem penglihatan manusia disaat otak memproses gambar yang berasal dari mata. Studi forensik dan biometrik yakni cara mengenali orang menggunakan visi komputer termasuk pengenalan wajah otomatis dan mengenali orang dengan tekstur atau iris mereka. Studi ini disejajarkan dengan ahli biologi dan psikolog yang terus mempelajari bagaimana sistem penglihatan manusia bekerja dan

bagaimana melihat dan mengenali objek seperti mengenali seseorang (Nixon & Aguado, 2019).

#### D. Pengolahan Citra Digital

Pengolahan citra digital adalah pemrosesan gambar dua dimensi menggunakan komputer dan mengacu pada pemrosesan setiap data dua dimensi. Citra digital merupakan sebuah larik yang berisi nilai nilai real maupun kompleks yang direpresentasikan dengan deretan bit tertentu. Suatu citra data didefinisikan sebagai fungsi  $f(x, y)$  berukuran  $M$  untuk baris dan  $N$  untuk kolom, dimana  $x$  dan  $y$  adalah koordinat spasial dan amplitude  $f$  di titik koordinat  $(x, y)$  dinamakan intensitas atau tingkat keabuan dari citra pada titik tersebut adalah citra digital, sebagaimana yang ditunjukkan pada gambar 2.1.



Gambar 2.1. Koordinat Citra Digital

Sebuah citra digital dapat diwakili oleh sebuah matriks dua dimensi  $f(x, y)$  yang terdiri dari  $M$  kolom dan  $N$  baris, dimana perpotongan antara kolom dan baris disebut piksel (*pixel = picture element*) atau elemen terkecil dari sebuah citra sebagaimana ditunjukkan pada gambar 2.2. Adapun

bentuk bentuk matriks  $f(x, y)$ , sebagaimana yang ditunjukkan persamaan (1) berikut ini.

$$f(x, y) \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0, N-1) \\ f(1,0) & f(1,1) & \dots & f(1, N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1,0) & f(M-1,1) & \dots & f(M-1, N-1) \end{bmatrix} \dots \dots \dots (1)$$

Suatu citra  $f(x, y)$  dalam fungsi matematis dapat dituliskan sebagai berikut:

$$\begin{aligned} 0 &\leq x \leq M-1 \\ 0 &\leq y \leq N-1 \\ 0 &\leq f(x, y) \leq G-1 \end{aligned}$$

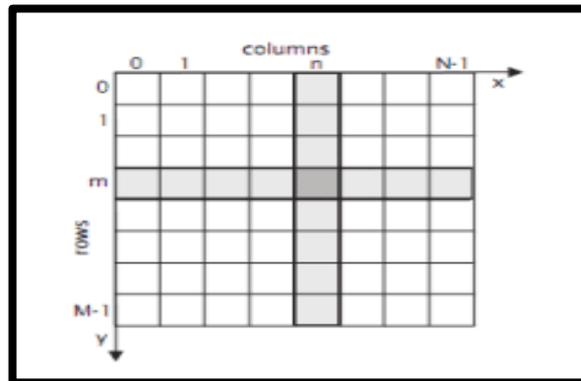
dimana:

$M$  = jumlah piksel baris (*row*) pada array citra  
 $N$  = jumlah piksel kolom (*column*) pada array citra  
 $G$  = Nilai skala keabuan (*grayscale*)

Besarnya nilai  $M$ ,  $N$  dan  $G$  pada umumnya merupakan perpangkatan dari dua.

$$M = 2^m; N = 2^n; G = 2^k \dots \dots \dots (2)$$

dimana nilai  $m$ ,  $n$  dan  $k$  adalah bilangan bulat positif. Interval  $(0, G)$  disebut skala keabuan (*grayscale*). Besar  $G$  tergantung pada proses digitalisasinya. Biasanya keabuan 0 (nol) menyatakan intensitas hitam dan 1 (satu) menyatakan intensitas putih. Untuk citra 8 bit, nilai  $G$  sama dengan  $2^8 = 256$  warna (derajat keabuan) (Putra, 2010) (Kusumanto & Tomponu, 2016).



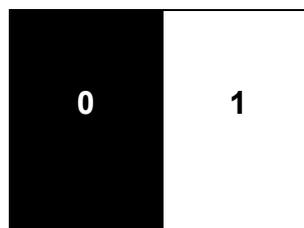
Gambar 2.2. Representasi citra digital dalam dua dimensi

### E. Jenis Citra Digital

Ada banyak cara untuk menyimpan citra digital di dalam memori, cara penyimpanan menentukan jenis citra digital yang terbentuk. Beberapa jenis citra digital yang sering digunakan adalah citra biner, citra skala keabuan, dan citra warna (T. Sutoyo, S.Si, Edy Mulyanto, S.Si., 2009).

#### 1. Citra Biner (*Monokrom*)

Banyaknya warna yakni dua, yaitu hitam dan putih, dibutuhkan 1 bit di memori untuk menyimpan kedua warna ini, dimana gradasi warnanya adalah bit 0 untuk warna hitam dan bit 1 untuk warna putih, seperti yang ditunjukkan pada gambar 2.3.



Gambar 2.3. Citra biner dua warna

#### 2. Citra skala keabuan (*grayscale*)

Banyaknya warna tergantung pada jumlah bit yang disediakan di memori untuk menampung kebutuhan warna. Citra dua bit mewakili 4 warna

dengan gradasi warna sebagaimana yang ditunjukkan oleh gambar 2.4, sedangkan untuk Citra tiga bit, mewakili 8 warna dengan gradasi warna sebagaimana terlihat pada gambar 2.5.

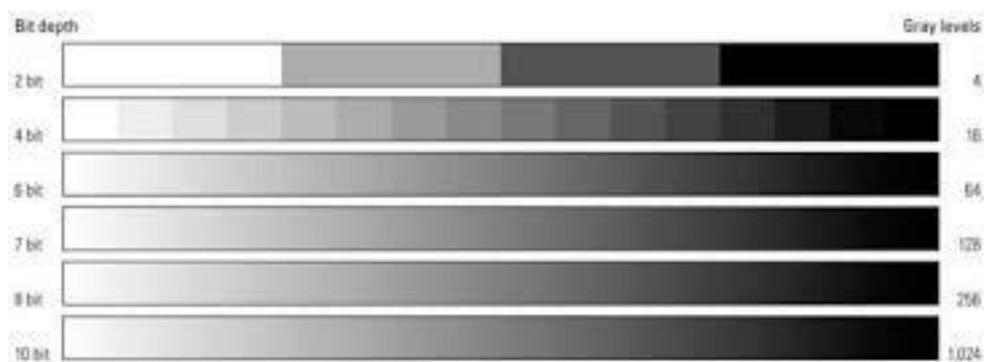


Gambar 2.4. Citra biner empat warna



Gambar 2.5. Citra biner delapan warna

Semakin besar jumlah bit warna yang disediakan di memori, maka semakin halus gradasi warna yang terbentuk. Gambar 2.6 menunjukkan perbandingan gradasi warna untuk jumlah bit mulai dari 2 bit, sampai dengan 10 bit.



Gambar 2.6. Perbandingan gradasi warna mulai 2 bit - 10 bit

### 3. Citra Warna (*True Color*)

Setiap piksel pada citra warna mewakili warna yang merupakan kombinasi dari tiga warna dasar (RGB = *Red, Green, Blue*). Setiap warna dasar menggunakan penyimpanan 8 *bit* = 1 *byte*, yang berarti setiap warna mempunyai gradasi sebanyak 255 warna. Berarti setiap piksel

mempunyai kombinasi warna sebanyak  $2^8 \cdot 2^8 \cdot 2^8 = 2^{24}$  16 juta warna lebih. Itulah sebabnya format ini dinamakan *true color* karena mempunyai jumlah warna yang cukup besar sehingga bisa dikatakan hampir mencakup semua warna di alam.

Penyimpanan citra warna di dalam memori berbeda dengan citra skala keabuan. Setiap piksel dari citra skala keabuan 256 gradasi warna diwakili oleh 1 *byte*. Sedangkan 1 piksel citra warna diwakili oleh 3 *byte*, dimana masing-masing *byte* merepresentasikan warna merah, hijau, dan biru, gambar 2.7 adalah contoh penyimpanan citra warna di dalam memori.

|  |   |   |
|--|---|---|
| <b>Yellow</b><br>R = 255<br>G = 255<br>B = 0 | <b>Orange</b><br>R = 255<br>G = 102<br>B = 0      | <b>Green</b><br>R = 0<br>G = 255<br>B = 0     |
| <b>Cyan</b><br>R = 0<br>G = 255<br>B = 255   | <b>Violet</b><br>R = 204<br>G = 102<br>B = 204    | <b>White</b><br>R = 255<br>G = 255<br>B = 255 |
| <b>Black</b><br>R = 0<br>G = 0<br>B = 0      | <b>Turquoise</b><br>R = 102<br>G = 255<br>B = 204 | <b>Brown</b><br>R = 153<br>G = 102<br>B = 51  |

Gambar 2.7. Penyimpanan citra warna di dalam memori

## F. Elemen Citra Digital

Berikut ini merupakan elemen-elemen yang terdapat pada citra digital (Putra, 2010):

1. Kecerahan warna (*brightness*)

Kecerahan merupakan intensitas cahaya yang dipancarkan piksel dari citra yang dapat ditangkap oleh sistem penglihatan. Kecerahan pada sebuah titik (piksel) di dalam citra merupakan intensitas rata-rata dari suatu area yang melingkupinya.

## 2. Kontras (*contrast*)

Kontras menyatakan sebaran terang dan gelap dalam sebuah citra. Pada citra yang baik, komposisi gelap dan terang tersebar secara merata.

## 3. Kontur (*contour*)

Kontur adalah keadaan yang ditimbulkan oleh perubahan intensitas pada piksel –piksel yang bertetangga. Karena adanya perubahan intensitas, maka mata mampu mendeteksi tepi-tepi objek di dalam citra.

## 4. Warna

Warna sebagai persepsi yang ditangkap sistem visual terhadap panjang gelombang cahaya yang ditentukan oleh objek.

## 5. Bentuk (*shape*)

Bentuk adalah property intrinsik dari objek 3 dimensi, dengan pengertian bahwa bentuk merupakan properti intrinsic utama untuk sistem visual manusia.

## 6. Tekstur (*texture*)

Tekstur dicirikan sebagai distribusi spasial dari derajat keabuan di dalam sekumpulan piksel-piksel yang bertetangga. Tekstur adalah sifat-sifat atau karakteristik yang dimiliki oleh suatu daerah yang cukup besar sehingga secara alami sifat-sifat tadi dapat berulang dalam daerah

tersebut. Tekstur adalah keteraturan pola-pola tertentu yang terbentuk dari susunan piksel-piksel dalam citra digital. Informasi tekstur dapat digunakan untuk membedakan sifat-sifat permukaan suatu benda dalam citra yang berhubungan dengan kasar dan halus, juga sifat-sifat spesifik dari kekasaran dan kehalusan permukaan tadi, yang sama sekali terlepas dari warna permukaan tersebut.

### **G. Resolusi Citra Digital**

Resolusi citra adalah tingkat detail dari suatu citra, semakin tinggi resolusi sebuah citra digital, maka semakin tinggi tingkat detail dari citra tersebut. Resolusi merupakan jumlah piksel yang tersusun dalam sebuah gambar digital. Resolusi berbanding lurus dengan kualitas gambar, karena itu resolusi juga merupakan salah satu faktor penentu kualitas dari citra digital. Semakin tinggi resolusi, semakin baik kualitas gambar digital, dan sebaliknya jika resolusinya semakin rendah maka semakin rendah kualitas gambar.

Setiap jenis citra memiliki keunggulan masing-masing dalam hal resolusi, suatu citra dapat unggul untuk resolusi tertentu namun lemah untuk jenis resolusi yang lain. Terdapat dua jenis resolusi citra digital yang perlu diketahui (T. Sutoyo, S.Si, Edy Mulyanto, S.Si., 2009), yaitu:

#### **1. Resolusi Spasial**

Resolusi spasial ini merupakan ukuran halus (resolusi tinggi) atau kasar (resolusi rendah) pada pembagian kisi-kisi baris dan kolom pada saat sampling, resolusi ini dipakai untuk menentukan jumlah pixel per satuan

panjang. Biasanya satuan resolusi ini adalah dpi (dot per inchi), resolusi ini sangat berpengaruh pada detail dan perhitungan gambar.

## 2. Resolusi Kecemerlangan (Brightness)

Resolusi kecermerlangan (intensitas/ brightness) atau biasanya disebut dengan kedalaman bit/kedalaman warna (Bit Depth) adalah ukuran halus (resolusi tinggi) atau kasar (resolusi rendah) pembagian tingkat gradasi warna saat dilakukan kuantisasi. Bit Depth menentukan berapa banyak informasi warna yang tersedia untuk ditampilkan dalam setiap piksel. Semakin besar nilainya, semakin bagus kualitas gambar yang dihasilkan dan tentu ukuran juga semakin besar.

Pada sistem pengenalan wajah yang telah berkembang sejak tahun 1960-an, untuk mengidentifikasi seseorang dari gambar digital atau video, citra atau gambar diklasifikasi menjadi dua bagian yaitu gambar dengan resolusi tinggi (halus) dan gambar dengan resolusi rendah (kasar). Adapun ukuran gambar yang dikategorikan sebagai gambar dengan resolusi rendah yakni 64 x 64 pixels, 32 x 32 pixels dan 16 x 16 pixels. Gambar dengan kategori beresolusi tinggi dimulai dari ukuran 128 x 128 pixels (Engineering & Gables, 2017).

## H. Video Digital

Video adalah sekumpulan citra yang direkam atau diakuisisi selama satu satuan waktu tertentu. Citra di dalam video lebih dikenal dengan istilah frame atau frame citra. Kecepatan standar perekaman dan visualisasi video disesuaikan dengan kecepatan mata manusia dalam memahami dan menganalisis informasi yang terkandung di dalam setiap frame. Pada

kamera video standar Eropa kecepataannya adalah 25 frame per detik/frame per second (fps), sedangkan untuk standar Amerika adalah 30 fps. Jika sekumpulan citra ditampilkan ke layar monitor dengan kecepatan tampil antara 25 hingga 30 fps maka mata manusia dapat mengamati dengan baik informasi dan pergerakan objek di dalam setiap frame video secara kontinu. Bila tampilan frame video ini diperlambat menjadi di bawah 15 fps maka akan tampak pergerakan objek yang tidak kontinyu (terputus-putus), sebaliknya bila tampilan frame video ini dipercepat di atas 40 fps maka pergerakan objek akan semakin cepat dan mata manusia semakin sulit untuk menganalisis informasi pergerakan objek (Madenda, 2015).

Video terbagi dalam dua kategori yaitu video analog dan video digital. Video analog mengkodekan informasi gambar dengan memvariasikan voltase dan/atau frekuensi dari sinyal. Seluruh sistem sebelum video digital dapat dikategorikan sebagai video analog. Sementara video digital dapat disebut array 3 dimensi dari pixel berwarna, 2 dimensi melayani arah spesial dari gambar bergerak (horizontal dan vertikal) dan satu dimensinya akan mempresentasikan domain waktu (Wibowo et al., 2017).

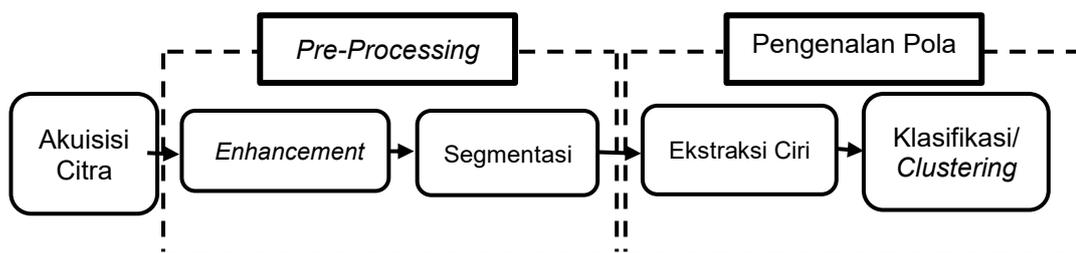
### **I. Pengenalan Pola**

Pengenalan pola merupakan cabang dari kecerdasan buatan, untuk menyelesaikan masalah yang kompleks. Pengenalan pola memberikan solusi bagi masalah yang berkaitan dengan pengenalan atau klasifikasi, seperti pengenalan wajah, pengenalan suara, klasifikasi teks dokumen, pengenalan pola DNA & sidik jari, klasifikasi karakter tulisan tangan,

pengenalan pola batik, pengenalan pola daging sapi, daging babi maupun daging ayam dan diagnosis medis.

Pengenalan pola adalah salah satu bidang dalam pembelajaran mesin (*machine learning*) yang menitikberatkan pada metode klasifikasi objek ke dalam kelas-kelas tertentu untuk menyelesaikan masalah tertentu. Aplikasi pengenalan pola bertujuan untuk melakukan proses pengenalan terhadap suatu objek ke dalam salah satu kelas tertentu berdasarkan pola yang dimilikinya. Di dalam mengenali sebuah pola, komputer memerlukan sebuah algoritma untuk menentukan tingkat kesamaan antara pola uji dan pola yang ada pada basis data.

Bidang pengenalan pola telah banyak dikembangkan oleh para peneliti, sehingga menghasilkan variasi algoritma pengenalan pola. Semua algoritma tersebut terdiri dari tiga elemen dasar, yaitu, persepsi data, ekstraksi ciri dan klasifikasi. Banyak teknik yang berbeda dalam mengimplementasikan tiga elemen dasar ini. Algoritma pengenalan pola secara garis besar bisa digambarkan sebagai berikut:



Gambar 2.8. Algoritma pengenalan pola

Berikut ini merupakan penjelasan dari gambar 2.8 tentang algoritma pengenalan pola:

1. **Akuisisi citra** adalah proses pengubahan citra analog menjadi citra digital yang diambil dari lingkungan atau dunia nyata menggunakan beberapa alat berikut, kamera digital, *webcam*, *smartphone*, *scanner*, dan lain-lain, agar bisa dilanjutkan ke tahap *pre-processing*. Terdapat beberapa hal penting yang sangat mempengaruhi citra digital yang dihasilkan saat dilakukan proses akuisisi, yaitu resolusi alat yang digunakan, jarak dan sudut pandang pengambilan citra, faktor pencahayaan, perbesaran dan pengecilan, objek atau kamera dalam keadaan bergerak atau tidak, serta format citra yang dihasilkan.
2. ***Pre-processing*** adalah tahap pemrosesan data (dalam hal ini citra digital) agar data bisa dan layak digunakan untuk tahap berikutnya. Hal ini dilakukan karena hasil citra digital dari proses akuisisi biasanya memiliki beberapa masalah, misalnya terjadinya *noise* atau adanya objek-objek pengganggu. Beberapa masalah tersebut disebabkan oleh kurang akuratnya sensor atau transduser yang digunakan saat proses akuisisi. Selain itu, tujuan *pre-processing* adalah untuk membuat citra digital agar sesuai dengan kebutuhan ekstraksi fiturnya.
3. ***Image enhancement*** (perbaikan kualitas citra) merupakan bagian dari tahap *preprocessing* yang bertujuan untuk meningkatkan kualitas citra. Sebuah citra dikatakan mempunyai kualitas baik jika pada citra tersebut dilakukan segmentasi yang hasilnya dapat tersegmentasi dengan baik. Beberapa metode *image enhancement* yang biasa digunakan adalah

*intensity adjustment, contrast stretching, filtering (median filter, low pass filter, high pass filter dan lain-lain).*

4. **Segmentasi** merupakan bagian dari tahap *pre-processing* yang bertujuan untuk memisahkan objek tertentu yang dikehendaki (*foreground*) dengan objek lain yang tidak dikehendaki (*background*) *output* (masukan) segmentasi biasanya berupa citra biner, di mana *foreground* diberi simbol 1 dan *background* diberi simbol 0. Beberapa metode segmentasi yang biasa digunakan adalah *Thresholding*, *Multithresholding*, deteksi tepi, K-Means Clustering, Filter Gabor, Fuzzy C-Means Clustering, Watershed, Transformasi Hough dan lain- lain.

a) **Ekstraksi ciri** merupakan tahap pengambilan ciri objek yang bisa menjadi bahan pembeda dari objek-objek lainnya.

b) **Klasifikasi / clustering** adalah algoritma yang mampu mengklasifikasikan / meng-cluster objek berdasarkan pada karakteristik ciri-ciri yang diberikan. *Keluaran* dari algoritma ini adalah objek yang akan dikelompokkan ke dalam kelas yang mempunyai kesamaan ciri, misalnya sebuah objek masuk dalam kelas apel Amerika atau apel Malang maupun sebuah objek masuk dalam kelas daging sapi atau daging babi atau daging ayam. Jadi, dapat dikatakan bahwa klasifikasi merupakan pengelompokan data (parameter hasil dari ekstraksi ciri) berdasarkan pada kesamaan ciri yang dimiliki oleh data tersebut. Algoritma klasifikasi yang biasa digunakan adalah *K-Nearest Neighbor*, *Naive Bayes*, Jaringan

Syaraf Tiruan, *Support Vector Machine* (SVM), Decision Trees dan lain-lain.

- c) **Teknik pembelajaran *Supervised* dan *Unsupervised*** adalah Teknik pembelajaran pada algoritma pengenalan pola tergantung pada tipe label *output* yang digunakan. Jika label *output* data pembelajaran diketahui, maka disebut sebagai pembelajaran *supervised* dan pembelajaran seperti ini digunakan untuk keperluan klasifikasi. Jika label *output* data pembelajaran tidak diketahui, maka disebut sebagai pembelajaran *Unsupervised* dan pembelajaran seperti ini digunakan untuk keperluan *clustering* (Andono et al., 2017).

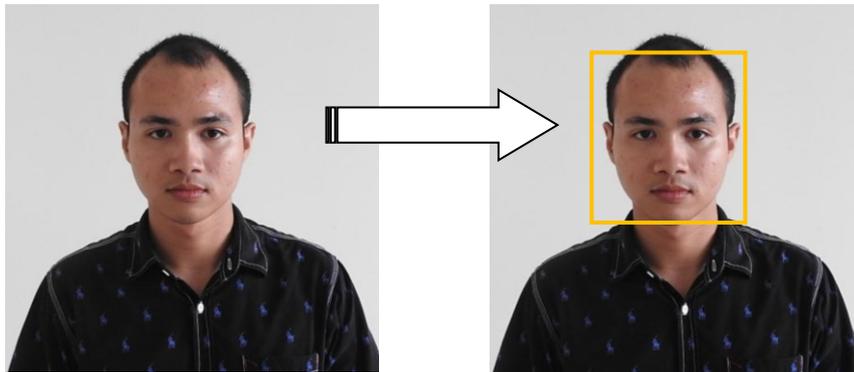
#### **J. Deteksi Wajah**

Deteksi wajah adalah teknologi komputer yang digunakan dalam berbagai aplikasi yang mengidentifikasi lokasi adanya wajah manusia dalam gambar digital. Deteksi wajah juga merupakan salah satu tahap pra-proses yang sangat penting di dalam sistem pengenalan wajah (*face recognition*) yang digunakan untuk sistem *biometric* (Viola & Jones, 2001). Deteksi wajah juga dapat digunakan untuk pencarian dan pengindeksan citra atau video yang di dalamnya terdapat wajah manusia dalam berbagai ukuran, posisi, dan latar belakang. Pemrosesan gambar ini bertujuan untuk mencari wajah dari gambar obyek bergerak yang telah di-capture, kemudian gambar tersebut diolah dengan memisahkan gambar dengan latar belakangnya, sehingga hanya bagian yang dianggap kulit yang ditampilkan sedangkan bagian yang bukan kulit akan dihitamkan.

Pemisahan gambar dengan latar belakang ini bertujuan untuk memudahkan proses pencarian wajah.

Deteksi wajah adalah menentukan keberadaan wajah dalam gambar dan, jika ada, mencari lokasi gambar dan bagian masing-masing wajah. Deteksi wajah merupakan hal yang mudah bagi manusia, namun sangat sulit untuk komputer (Yang et al., 2016). Tujuan utama dari deteksi wajah adalah mengetahui semua posisi wajah pada gambar dan memberi tanda dengan kotak persegi (Ma & Wang, 2019).

Deteksi wajah adalah langkah awal untuk melakukan identifikasi wajah, analisis wajah, estimasi pose wajah dan lainnya. Sebuah pendeteksi wajah yang ideal seharusnya mampu mengidentifikasi dan menemukan lokasi dan luas semua wajah yang ada di dalam sebuah gambar tanpa memperhatikan pose, skala, orientasi, umur, dan ekspresi (Li et al., 2004). Tingkat pemrosesan yang lebih tinggi dapat dilakukan hanya jika wajah ditandai secara akurat dalam gambar atau video dan landmark wajah secara akurat berada. Proses pendeteksi wajah ini bekerja dengan cara memeriksa citra yang dimasukkan, apakah memiliki citra wajah atau tidak, jika memiliki, maka akan dilakukan pemisahan dengan cara memotong citra wajah dari latar belakang citra yang dimasukkan. Berikut merupakan gambar pendeteksian wajah.



Gambar 2. 9. Ilustrasi Deteksi Wajah

### K. Pengenalan Wajah

Pengenalan Wajah adalah salah satu teknik *biometric* yang memungkinkan komputer untuk mengenali wajah manusia yang berorientasi pada wajah yang telah diterapkan untuk banyak aplikasi dalam bidang keamanan, antara lain *Access security system*, *Authentication system*, hingga sebagai alat bantu dalam pelacakan pelaku kriminal. Pengenalan ini dapat dibagi menjadi dua bagian yaitu: “dikenali” dan “tidak dikenali”, setelah dilakukan dengan pola yang sebelumnya disimpan di dalam *database*. Sistem ini juga harus bisa mengenali objek bukan wajah.

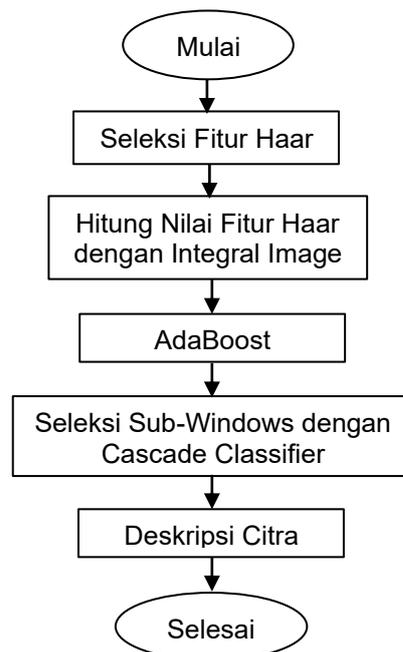
Secara umum sistem pengenalan citra wajah dibagi menjadi dua jenis, yaitu sistem *feature-based* dan sistem *image-based*. Pada sistem pertama digunakan ciri yang dieksekusi dari komponen citra wajah seperti mata, hidung, mulut, dan lain-lain yang kemudian dimodelkan secara geometris hubungan antara ciri-ciri tersebut. Sedangkan pada sistem kedua menggunakan informasi dari pixel citra yang kemudian direpresentasikan dalam metode tertentu, yang kemudian digunakan untuk klasifikasi identitas citra (Fatta, 2009).

### **L. Oklusi Pengenalan Wajah**

Oklusi berasal dari kata *occlusion*, yang terdiri dari dua kata yakni *oc* yang berarti ke atas (*up*) dan *clusion* yang berarti menutup (*closing*). Jadi oklusi pada pengenalan wajah merupakan wajah – wajah yang tertangkap oleh kamera yang dapat tertutup sebagian dari wajahnya antara lain oleh aksesoris, janggut, atau kumis. Oklusi pada pengenalan wajah adalah tantangan untuk setiap peneliti bagaimana cara mengurangi halangan pada proses pengenalan wajah yang akan dilakukan oleh sistem (Yang et al., 2016).

### **M. Metode Viola-Jones**

Metode Viola-Jones digunakan untuk mendeteksi wajah atau citra. Proses pendeteksian wajah dilakukan dengan mengklasifikasikan sebuah gambar setelah sebelumnya sebuah pengklasifikasian dibentuk dari sebuah data (Viola & Jones, 2001). Klasifikasi citra dilakukan berdasarkan nilai dari sebuah fitur. Penggunaan fitur dilakukan karena pemrosesan fitur berlangsung lebih cepat dibandingkan pemrosesan citra per piksel. Metode Viola-Jones berpusat pada penghitungan setiap piksel sehingga perhitungannya harus benar-benar terperinci. Flowchart perhitungan metode viola-jones dapat dilihat pada Gambar berikut.



Gambar 2. 10. Flowchart perhitungan metode viola-jones

Pendekatan untuk mendeteksi objek di dalam gambar adalah menggabungkan empat konsep utama (Viola & Jones, 2001), yaitu:

1. Fitur segi empat sederhana, yang disebut *Haar Feature*.
2. *Integral Image* untuk pendeteksian fitur secara cepat.
3. Metoda Machine *AdaBoost Learning*.
4. *Classifier of Cascade*, klasifikasi bertingkat untuk menggabungkan banyak fitur secara efisien.

## N. Support Vector Machine

*Support Vector Machine* (SVM) adalah algoritma yang dapat digunakan untuk klasifikasi pola (*pattern*) atau didefinisikan sebagai model keputusan terawasi yang terhubung dengan beberapa algoritma pembelajaran lainnya untuk melakukan tugas *machine learning* dan teori-

teori statistik. SVM banyak digunakan untuk menganalisis data, terutama di bidang pengolahan gambar, SVM berfungsi sebagai *model classifier* atau sebagai model regresi untuk memberikan informasi yang bermakna dengan menganalisis data masukan (Islam et al., 2018).

### **O. Histogram of Oriented Gradient**

*Histogram of Oriented Gradient* (HOG) digunakan untuk mengekstraksi fitur pada objek gambar dengan menggunakan objek manusia (Endra et al., 2018). Berdasarkan langkahnya, proses awal pada metode HOG adalah mengkonversi citra menjadi *grayscale*, yang kemudian dilanjutkan dengan menghitung nilai gradien setiap piksel. Setelah mendapatkan nilai gradien, maka proses selanjutnya yaitu menentukan jumlah bin orientasi yang akan digunakan dalam pembuatan histogram.

Proses ini disebut *spatial orientation binning*. Namun sebelumnya pada proses *gradient compute* gambar pelatihan dibagi menjadi beberapa *cell* dan dikelompokkan menjadi ukuran lebih besar yang dinamakan *block*. Sedangkan pada proses normalisasi *block* digunakan perhitungan geometri R-HOG. Proses ini dilakukan karena terdapat *block* yang saling tumpang tindih. Berbeda dengan proses pembuatan histogram citra yang menggunakan nilai – nilai intensitas piksel dari suatu citra atau bagian tertentu dari citra untuk pembuatan histogramnya, berikut adalah tahapan dalam HOG (Dalal, 2005):

1. Konversi Citra atau Normalisasi Warna
2. *Gradient Compute*

3. *Spatial Orientation Binning*

4. *Normalization Block*

5. *Detector Windows*

#### **P. Contrast Limited Adaptive Histogram Equalization**

*Contrast Limited Adaptive Histogram Equalization* (CLAHE) adalah generalisasi dari metode *Adaptive Histogram Equalization* (AHE). Metode ini mampu menghasilkan citra yang lebih baik dibandingkan dengan citra asli yang belum diproses (Sebatubun, 2016). Berbeda dengan AHE yang bekerja secara global, algoritma CLAHE membagi citra ke dalam area-area yang lebih kecil dan menerapkan AHE untuk masing-masing area tersebut. Algoritma CLAHE dapat dijelaskan sebagai berikut (Ramya & Subha Rani, 2012):

1. Citra asli dibagi menjadi sub-citra yang berukuran  $M \times N$ .
2. Menghitung histogram dari setiap sub-citra.
3. *Clipped histogram* dari setiap citra.

Jumlah piksel pada sub-citra akan didistribusi pada masing-masing derajat keabuan. Rata-rata jumlah piksel untuk setiap derajat keabuan dirumuskan pada persamaan (4).

$$N_{avg} = \frac{N_{CR-Xp} * N_{CR-Yp}}{N_{gray}} \dots \dots \dots (3)$$

Dimana

$N_{avg}$  = Rata – rata jumlah piksel

$N_{gray}$  = Jumlah nilai derajat keabuan pada sub – citra

$N_{CR-Xp}$  = Jumlah piksel dalam dimensi X dari sub – citra

$N_{CR-Yp}$  = Jumlah piksel dalam dimensi Y dari sub – citra

Berdasarkan pada persamaan (3), maka *cliplimit* dapat dihitung menggunakan persamaan (4):

$$N_{CL} = N_{CLIP} * N_{avg} \dots \dots \dots (4)$$

Dimana

$N_{CL} = \text{Cliplimit}$

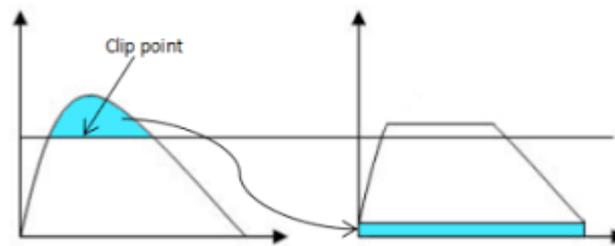
$N_{CLIP}$  = Nilai maksimum rata-rata piksel setiap nilai derajat keabuan dari sub-citra.

$N_{avg}$  = Rata – rata jumlah piksel

Permasalahan peningkatan kontras yang berlebihan pada AHE dapat diatasi dengan menggunakan CLAHE, yaitu memberikan nilai batas pada histogram. Nilai batas ini disebut dengan *cliplimit* yang menyatakan batas maksimum tinggi suatu histogram (Reza, 2004). Cara menghitung *cliplimit* suatu histogram dapat didefinisikan dengan persamaan (5).

$$\beta = \frac{M}{N} \left( 1 + \frac{a}{100} (s_{max} - 1) \right) \dots \dots \dots (5)$$

Variabel M menyatakan luas region size, N menyatakan nilai grayscale (256), merupakan *clip factor* menyatakan penambahan batas *limit* suatu histogram yang bernilai antara 0 sampai dengan 100. Histogram di atas nilai *cliplimit* dianggap kelebihan (*excess*) piksel yang akan didistribusikan kepada area sekitar di bawah *cliplimit*, sehingga histogram merata. Ilustrasi distribusi *excess* piksel dapat dilihat pada gambar 2.11.



Gambar 2. 11. Ilustrasi distribusi excess piksel di bawah *cliplimit*

### Q. Convolutional Neural Network

*Convolutional network* atau yang dikenal dengan *Convolutional Neural Network* (CNN) adalah tipe khusus dari *neural network* untuk memproses data yang mempunyai topologi jala atau *grid-like topology*. Pemberian nama *convolutional neural network* mengindikasikan bahwa jaringan tersebut menggunakan operasi matematika yang disebut konvolusi. Konvolusi sendiri adalah sebuah operasi linear. Jadi *convolutional network* adalah *neural network* yang menggunakan konvolusi minimal pada salah satu lapisannya (Lecun et al., 2015). *Convolutional neural network* merupakan *special case* dari *Artificial Neural Network* (ANN) yang saat ini diklaim sebagai model terbaik untuk memecahkan masalah *object recognition* dan *detection*.

CNN termasuk dalam jenis *Deep Neural Network* karena kedalaman jaringan yang tinggi dan banyak diaplikasikan pada data citra. Secara teknis, *convolutional network* adalah arsitektur yang bisa dilatih dan terdiri dari beberapa tahap. Masukan dan keluaran dari masing-masing tahap adalah beberapa *array* yang disebut *feature map* atau peta fitur. *Output* dari masing-masing tahap adalah *feature map* hasil pengolahan dari semua lokasi pada masukan. Tujuan dilakukannya konvolusi pada data citra

adalah untuk mengekstraksi fitur dari citra masukan. Secara umum operasi konvolusi dapat ditulis dengan persamaan (6) berikut ini.

$$s(t) = (x * w)(t) \dots\dots\dots (6)$$

Pada fungsi  $s(t)$  menghasilkan output tunggal yaitu *Feature Map*, argumen pertama berupa *masukan* yang merupakan  $x$  dan argumen kedua yang merupakan  $w$  sebagai *kernel* atau *filter*. Jika melihat masukan sebagai citra dua dimensi, maka  $(t)$  bisa diasumsikan sebagai sebuah piksel dan menggantinya dengan  $i$  dan  $j$ . Oleh karena itu, untuk operasi konvolusi dengan lebih dari satu dimensi dapat digunakan persamaan berikut.

$$S_{(i,j)} = (K * I)_{(i,j)} = \sum \sum I_{(i-m,j-n)} K_{(m,n)} \dots\dots\dots (7)$$

Pada persamaan (7) merupakan perhitungan dalam operasi konvolusi dengan  $i$  dan  $j$  sebagai piksel dari sebuah citra. Perhitungannya bersifat komutatif dan muncul ketika  $K$  sebagai kernelnya serta  $I$  sebagai masukan dan kernel yang dapat dibalik relative terhadap masukan. Operasi konvolusi dapat dilihat sebagai perkalian matriks antara citra masukan dan filter dimana keluarannya dapat dihitung dengan *dot product*. Secara umum jenis *layer* (lapisan) pada CNN dibedakan menjadi dua yaitu:

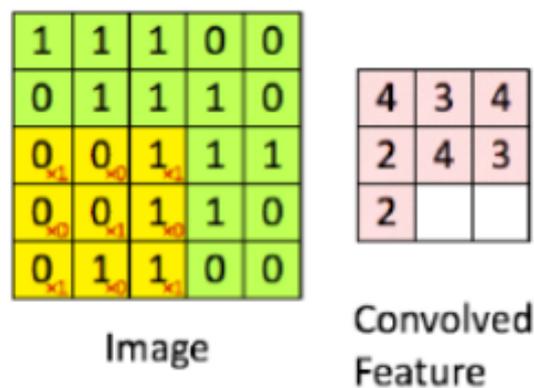
### **Q.1. Lapisan Ekstraksi Fitur Gambar**

Lapisan ini berada pada awal arsitektur tersusun atas beberapa lapisan dan setiap lapisan tersusun atas *neuron* yang terkoneksi pada daerah lokal (*local region*) lapisan sebelumnya. Lapisan ini menerima masukan gambar secara langsung dan memprosesnya hingga

menghasilkan keluaran berupa vektor untuk diolah pada Lapisan berikutnya. Lapisan ekstraksi fitur ini terdiri dari:

### 1. *Convolutional Layer*

Konvolusi adalah suatu istilah matematis yang berarti mengaplikasikan sebuah fungsi pada *output* fungsi lain secara berulang. Dalam pengolahan citra, konvolusi berarti mengaplikasikan sebuah kernel (kotak kuning) pada citra di semua offset yang memungkinkan seperti yang ditunjukkan pada gambar 2.12 Kotak hijau secara keseluruhan adalah citra yang akan di konvolusi. Kernel bergerak dari sudut kiri atas ke kanan bawah. Sehingga hasil konvolusi dari citra tersebut dapat dilihat pada gambar di sebelah kanannya.



Gambar 2.12. Operasi Konvolusi

Tujuan dilakukannya konvolusi pada data citra adalah untuk mengekstraksi fitur dari citra masukan, menurut (Liu et al., 2018) konvolusi adalah cara yang efisien untuk ekstraksi fitur. Konvolusi akan menghasilkan transformasi linear dari data masukan sesuai informasi spasial pada data. Bobot pada lapisan tersebut

menspesifikasikan kernel konvolusi yang digunakan, sehingga kernel konvolusi dapat dilatih berdasarkan masukan pada CNN.

## 2. *Rectified Linear Unit (ReLU)*

ReLU bertujuan untuk menjaga hasil citra proses konvolusi berada pada domain definit positif. ReLU merupakan salah satu fungsi aktivasi populer dalam *deep neural network*. Pada fungsi aktivasi dapat mengubah jumlah angka pembobotan dari masukan yang masuk ke dalam *neuron* buatan. Fungsi ini harus bersifat *non-linear* untuk pengkodean pola yang kompleks dari data. Aktivasi memiliki dua jenis yaitu *Sigmoid* dan *Tanh*. Persamaan (8) yang biasa digunakan dalam fungsi ini adalah:

$$f(x) = \max(x, 0) \dots\dots\dots (8)$$

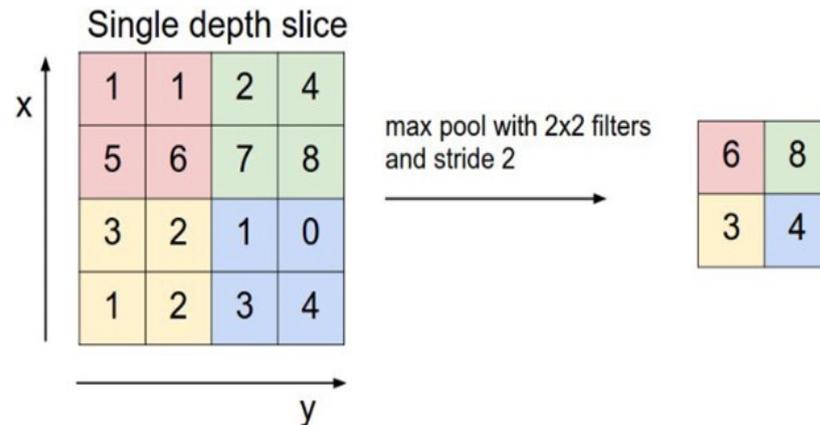
Dimana:

$x$  = nilai pada *feature map*

## 3. *Subsampling Layer* atau *Pooling Layer*

Proses ini bertujuan untuk mereduksi ukuran sebuah data citra. Dalam pengolahan citra, *subsampling* juga bertujuan untuk meningkatkan invariansi posisi dari fitur. Pada sebagian besar CNN, metode *subsampling* yang digunakan adalah *max pooling*. *Max pooling* membagi *output* dari *convolution layer* menjadi beberapa *grid* kecil lalu mengambil nilai maksimal dari setiap *grid* untuk menyusun matriks citra yang telah direduksi seperti yang ditunjukkan pada gambar 2.13. *Grid* yang berwarna merah, hijau, kuning dan biru merupakan kelompok *grid* yang akan dipilih nilai

maksimumnya. Sehingga hasil dari proses tersebut dapat dilihat pada kumpulan *grid* disebelah kanannya. Proses tersebut memastikan fitur yang didapatkan akan sama meskipun objek citra mengalami translasi (pergeseran).



Gambar 2.13. Operasi max pooling

## Q.2. Layer Klasifikasi

Layer ini menerima masukan dari hasil keluaran *layer* ekstraksi fitur gambar berupa vektor kemudian ditransformasikan seperti *Multi Neural Networks* dengan tambahan beberapa *hidden layer*. Hasil keluaran berupa skoring kelas untuk klasifikasi. Lapisan klasifikasi terdiri dari:

### 1. Flatten

Proses ini bertujuan untuk membentuk ulang fitur (*reshape feature map*) menjadi sebuah vektor agar bisa digunakan sebagai masukan dari *fully-connected layer*.

### 2. Fully-connected

Lapisan *Fully-connected* akan menghitung skor kelas. Seperti Jaringan Saraf biasa dan seperti namanya, setiap *neuron* dalam lapisan ini akan terhubung ke semua angka dalam volume.

### 3. *Softmax*

Fungsi *Softmax* menghitung probabilitas dari setiap kelas target atas semua kelas target yang memungkinkan dan akan membantu untuk menentukan kelas target untuk masukan yang diberikan. Keuntungan utama menggunakan *Softmax* adalah rentang probabilitas *output* dengan nilai 0 hingga 1, dan jumlah semua probabilitas akan sama dengan satu. Jika fungsi *softmax* digunakan untuk model multi-klasifikasi, maka akan mengembalikan peluang dari masing-masing kelas dan kelas target akan memiliki probabilitas tinggi. *Softmax* menggunakan eksponensial dari nilai masukan yang diberikan dan jumlah nilai eksponensial dari semua nilai dalam *input* (masukan). Maka rasio eksponensial dari nilai masukan dan jumlah nilai eksponensial adalah keluaran dari fungsi *softmax*.

## R. Arsitektur CNN

Uraian singkat tentang arsitektur CNN yang diperkenalkan para peneliti (Munir et al., 2019), diuraikan sebagai berikut.

### R.1. ResNet (*Deep Residual Network*)

*Deep Residual Network* atau yang biasa disebut sebagai ResNet merupakan salah satu arsitektur dari CNN. Arsitektur ini dibangun untuk mengatasi permasalahan pada pelatihan *Deep Learning* (DL),

karena pelatihan DL pada umumnya memakan cukup banyak waktu dan terbatas pada jumlah lapisan tertentu. Solusi permasalahan yang diusulkan oleh ResNet adalah dengan menerapkan *skip connection* atau *shortcut*. Kelebihan model ResNets dibandingkan dengan model arsitektur CNN yang lain adalah kinerja dari model ini tidak menurun walaupun arsitekturnya semakin dalam. Selain itu, perhitungan komputasi yang dilakukan lebih ringan dan kemampuan untuk melatih jaringan yang lebih baik. Model ResNet diimplementasikan dengan melakukan *skip connection* pada dua sampai tiga layer yang mengandung ReLU dan *batch normalization* di antara arsitekturnya (He et al., 2016).

## **R.2. MobileNet**

MobileNet merupakan salah satu arsitektur dari *convolutional neural network* (CNN). Para peneliti dari google membuat arsitektur tersebut untuk kebutuhan mobile. Perbedaan mendasar dari mobilenet dengan arsitektur CNN lainnya adalah pada penggunaan layer konvolusi dengan ketebalan filter yang sesuai dengan masukan gambar. MobileNet juga digunakan untuk konvolusi *depthwise* dan *pointwise* (Howard et al., 2017).

## **R.3. ArcFace**

*Additive Angular Margin Loss* (Arcface) merupakan model yang digunakan untuk meningkatkan kekuatan diskriminatif dari model pengenalan wajah dan untuk menstabilkan proses latih. *Loss function*

yang umum digunakan dalam proses klasifikasi adalah *Softmax Loss*, adapun persamaan yang digunakan adalah berikut:

$$L_{Softmax} = -\frac{1}{N} \sum_{i=1}^N \log \frac{e^{\frac{w_j^T x_i + b_j}{y_i}}}{\sum_{j=1}^n e^{\frac{w_j^T x_i + b_j}{y_i}}}, \dots \dots \dots (9)$$

Dimana  $x_i \in R^d$  merupakan fitur dari sampel ke- $i$  yang berada pada kelas  $y_i$  dan  $d$  merupakan dimensi *embedding* yang berukuran 512.  $W_j \in R^d$  adalah bobot dari kelas ke- $j$ , dan  $b_j \in R^n$  adalah bias dari kelas ke- $j$ . Kemudian  $N$  adalah jumlah sampel yang akan digunakan, dan  $n$  adalah jumlah kelas. Untuk mengoptimalkan *Softmax Loss* yang dapat memberikan tingkat kemiripan fitur yang tinggi antara sampel yang memiliki kelas yang sama sekaligus meningkatkan perbedaan fitur antara kelas yang berbeda. Adapun persamaan yang dari metode ArcFace (Deng, 2019) yakni.

$$L_{ArcFace} = -\frac{1}{N} \sum_{i=1}^N \log \frac{e^{s(\cos(\theta_{y_i} + m))}}{e^{s(\cos(\theta_{y_i} + m))} + \sum_{j=1, j \neq y_i}^n e^{s \cos \theta_j}} \dots \dots \dots (10)$$

$\cos \theta_j$  merupakan hasil dari  $\frac{W_j}{\|W_j\|} \otimes \frac{x_i}{\|x_i\|}$ .  $\theta_{y_i}$  didapatkan dari  $\arccos \arccos (\cos \theta_{j=y_i})$ .  $m$  merupakan *angular margin penalty* untuk  $\theta_{y_i}$ . Setelah  $\cos (\theta_{y_i} + m)$  dan  $\cos \theta_j$  dihitung, hasilnya kemudian akan di re-scale sebesar  $s$ .

### S. Cosine Similarity

*Cosine similarity* (CosSim) merupakan metode yang digunakan untuk menghitung tingkat kesamaan dari dua buah vektor. Persamaan antara dua vektor A dan B dapat dilihat pada persamaan (11). Cosine

Similarity memiliki sifat khusus yang membuatnya cocok untuk metrik, ukuran kemiripan yang dihasilkan selalu dalam kisaran  $-1$  dan  $+1$ , hal ini memungkinkan fungsi tujuan menjadi sederhana dan efektif (Nguyen & Bai, 2011).

$$\text{CosSim}(x, y) = \frac{A^T B}{\|A\| \|B\|} \dots \dots \dots (11)$$

### T. Evaluasi Sistem

Evaluasi kinerja pada suatu sistem klasifikasi merupakan sesuatu yang penting untuk mengetahui seberapa baik sistem yang dibangun dalam mengklasifikasikan data (Wijaya & Tjiharjadi, 2009). Berdasarkan jumlah kelas pada keluaran hasil sistem klasifikasi, jenis klasifikasi dikelompokkan menjadi empat jenis yaitu klasifikasi biner, multi-case, multi-label, dan hierarchical (Sokolova & Lapalme, 2009). Pada pengukuran kinerja sistem klasifikasi umumnya digunakan *Confusion Matrix* untuk membandingkan hasil klasifikasi oleh sistem dengan hasil klasifikasi sesungguhnya. Pada *Confusion Matrix*, dikenal empat istilah berupa *True Positive* (TP), *True Negative* (TN), *False Positive* (FP) dan *False Negative* (FN) untuk mempresentasikan hasil perbandingan klasifikasi sistem dengan klasifikasi sesungguhnya.

Pasangan gambar wajah yang dikenali dengan benar, memiliki identitas yang sama didefinisikan pada persamaan (12). Pada pasangan gambar dengan identitas berbeda yang diidentifikasi memiliki wajah yang sama didefinisikan pada persamaan (13). Nilai threshold  $d$  digunakan sebagai syarat untuk menentukan apakah kedua embedding memiliki

identitas yang sama atau berbeda. Jika nilai kemiripan yang dihasilkan lebih besar dari *threshold*, maka kedua *embedding* tersebut dikatakan memiliki identitas yang sama, begitu pula sebaliknya. Untuk Nilai True Positif Rate TPR(d) dan False Positive Rate FPR(d) dihitung dengan menggunakan persamaan (15) dan persamaan (16) (Champanand, 2015).

$$TA(d) = \{(a, b) \in P_{same}, \text{with } CosSim(s_a, t_b) \geq d\} \dots\dots\dots (12)$$

$$FA(d) = \{(a, b) \in P_{diff}, \text{with } CosSim(s_a, t_b) \geq d\} \dots\dots\dots (13)$$

$$TPR(d) = \frac{TA(d)}{P_{same}} \dots\dots\dots (14)$$

$$FPR(d) = \frac{FA(d)}{P_{diff}} \dots\dots\dots (15)$$

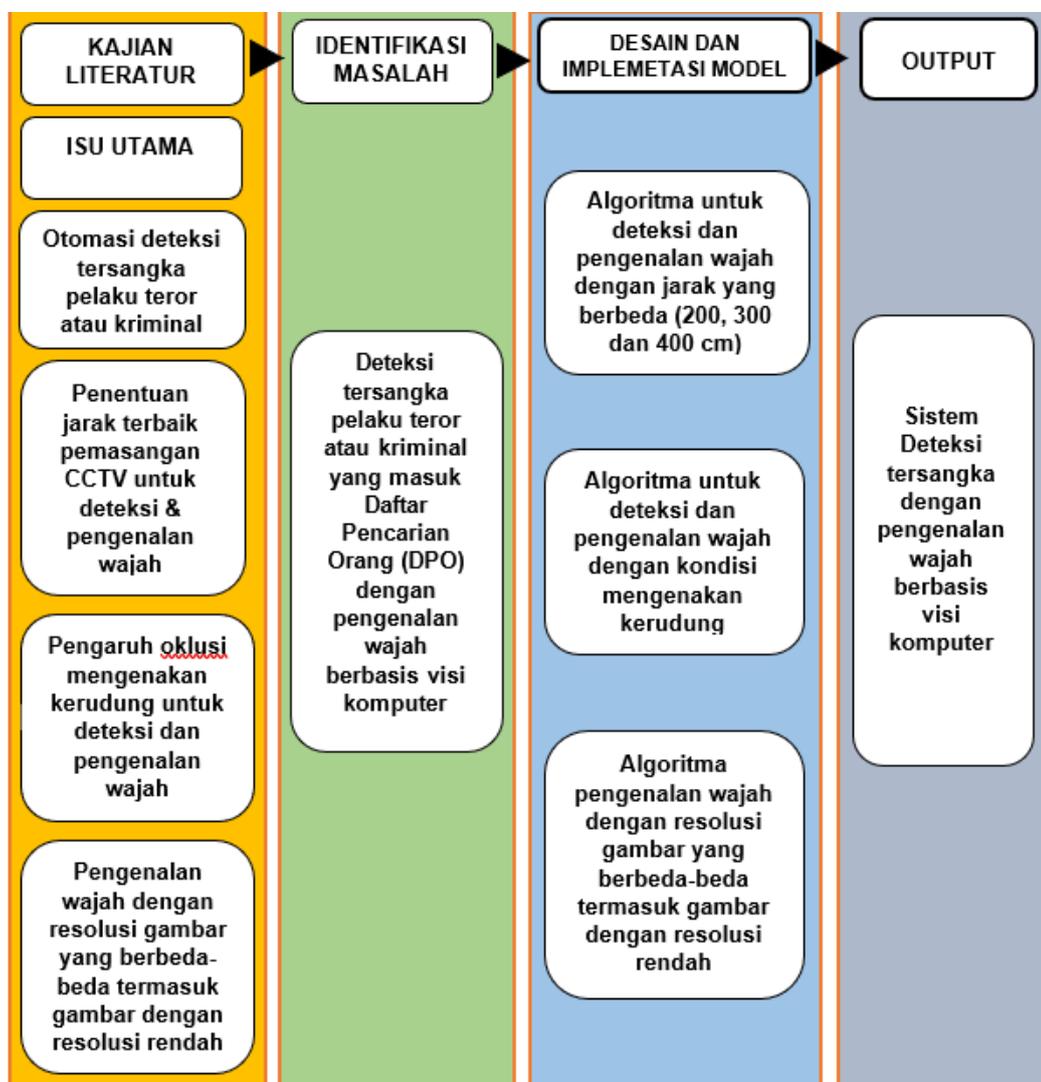
$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \dots\dots\dots (16)$$

## U. Kerangka Pikir Penelitian

Mengacu pada latar belakang dan rumusan masalah serta tujuan penelitian, kemudian berdasarkan juga hasil penelitian terdahulu dan kajian literatur maka disusunlah kerangka pikir penelitian sebagaimana yang ditampilkan pada gambar 2.18.

Pada diagram tersebut menggambarkan tentang alur sistem yang dibuat dan dikelompokkan menjadi empat tahapan yaitu kajian literatur dengan konten tentang otomasi deteksi tersangka pelaku teror atau kriminal yang masuk daftar pencarian orang. Pada bagian ini terdapat tiga isu utama yang dibahas yakni penentuan jarak terbaik untuk pemasangan CCTV dalam melakukan deteksi dan pengenalan wajah, lalu pengaruh oklusi mengenakan kerudung dan pengenalan wajah dengan resolusi gambar

yang berbeda-beda termasuk resolusi rendah. Selanjutnya identifikasi masalah dalam melakukan deteksi tersangka pelaku teror atau kriminal yang masuk daftar pencarian orang, dilanjutkan dengan desain dan implementasi model yang mengacu pada hasil identifikasi masalah, kemudian tahapan terakhir adalah output berupa sistem deteksi tersangka dengan pengenalan wajah berbasis visi komputer.



Gambar 2.18. Kerangka Pikir Penelitian