

SKRIPSI

**SISTEM REKOGNISI WAJAH DENGAN OKLUSI
PENGUNAAN MASKER SECARA REAL TIME**

Disusun dan diajukan oleh:

**NURINA RAHAYU
D121171003**



**PROGRAM STUDI SARJANA TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS HASANUDDIN
GOWA
2024**



Optimized using
trial version
www.balesio.com

LEMBAR PENGESAHAN SKRIPSI

SISTEM REKOGNISI WAJAH DENGAN OKLUSI PENGGUNAAN MASKER SECARA REAL TIME

Disusun dan diajukan oleh

Nurina Rahayu
D121171003

Telah dipertahankan di hadapan Panitia Ujian yang dibentuk dalam rangka Penyelesaian
Studi Program Sarjana Program Studi Teknik Informatika
Fakultas Teknik Universitas Hasanuddin
Pada tanggal 15 Maret 2024
dan dinyatakan telah memenuhi syarat kelulusan

Menyetujui,

Pembimbing Utama,

Pembimbing Pendamping,

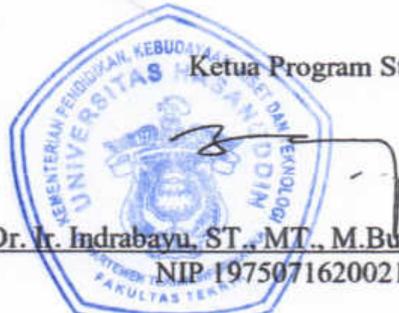


Prof. Dr. Ir. Indrabayu, ST., MT.,
M.Bus.Sys., IPM, ASEAN. Eng.
NIP 197507162002121004



Elly Warni, S.T., M.T.
NIP 198202162008122001

Ketua Program Studi,



Prof. Dr. Ir. Indrabayu, ST., MT., M.Bus.Sys., IPM, ASEAN. Eng.
NIP 197507162002121004



PERNYATAAN KEASLIAN

Yang bertanda tangan dibawah ini ;

Nama : Nurina Rahayu
NIM : D121171003
Program Studi : Teknik Informatika
Jenjang : S1

Menyatakan dengan ini bahwa karya tulisan saya berjudul

Sistem Rekognisi Wajah Dengan Oklusi Penggunaan Masker Secara Real Time

Adalah karya tulisan saya sendiri dan bukan merupakan pengambilan alihan tulisan orang lain dan bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri.

Semua informasi yang ditulis dalam skripsi yang berasal dari penulis lain telah diberi penghargaan, yakni dengan mengutip sumber dan tahun penerbitannya. Oleh karena itu semua tulisan dalam skripsi ini sepenuhnya menjadi tanggung jawab penulis. Apabila ada pihak manapun yang merasa ada kesamaan judul dan atau hasil temuan dalam skripsi ini, maka penulis siap untuk diklarifikasi dan mempertanggungjawabkan segala resiko.

Segala data dan informasi yang diperoleh selama proses pembuatan skripsi, yang akan dipublikasi oleh Penulis di masa depan harus mendapat persetujuan dari Dosen Pembimbing.

Apabila dikemudian hari terbukti atau dapat dibuktikan bahwa sebagian atau keseluruhan isi skripsi ini hasil karya orang lain, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Gowa, 15 Maret 2024

Yang Menyatakan



Nurina Rahayu



ABSTRAK

NURINA RAHAYU. *Sistem Rekognisi Wajah Dengan Oklusi Penggunaan Masker Secara Real Time* (dibimbing oleh Indrabayu dan Elly Warni)

Penggunaan masker saat ini yang dianggap wajar menyebabkan sulitnya mengenali wajah seseorang di tempat umum sehingga keadaan ini dimanfaatkan oleh suspek untuk berkamuflase agar tidak dikenali identitasnya apalagi masih menggunakan cara konvensional dengan mencocokkan data suspek dengan data yang telah disimpan di *database* sehingga membutuhkan waktu yang cukup lama dan tidak efisien.

Tujuan dari penelitian ini adalah untuk mengetahui bagaimana sistem merekognisi wajah dengan oklusi menggunakan masker secara real time dan bagaimana sistem ini bekerja baik dengan oklusi wajah maupun tanpa oklusi wajah.

Dalam membuat sistem rekognisi wajah, peneliti menggunakan Faster R-CNN (Region Based Convolutional Neural Networks) sebagai sistem pendeteksi dan klasifikasi. Apabila sistem mendeteksi wajah, baik dengan maupun tanpa oklusi, maka output yang diberikan berupa hasil klasifikasi sesuai dengan wajah yang terklasifikasi. Analisis kinerja sistem menggunakan Multi-Class Confusion Matrix dengan ukuran matriks 6x6 sesuai dengan jumlah kelas dalam sistem. Terdapat dua jenis pengujian, yaitu pengujian 1 dengan 345 gambar dan pengujian 2 dilakukan secara real time di ruangan yang telah dikondisikan.

Pengujian 1 menghasilkan nilai akurasi sebesar 100%. Pengujian 2 dengan dua skenario berdasarkan dua jarak berbeda yaitu 3 m dan 2 m dengan tujuh kondisi oklusi wajah yaitu tanpa oklusi, oklusi masker KF94 hitam, oklusi masker KF94 putih, oklusi masker duckbill hitam, oklusi masker duckbill putih, oklusi masker reguler hitam, dan oklusi masker reguler putih menghasilkan akurasi berturut-turut yaitu 97,64%, 95,56%, 87,18%, 99,13%, 95,37%, 94,51% dan 97,31% untuk jarak 3 m dan 99,8%, 100%, 99,3%, 99,49%, 99,81%, 99,63% dan 100% untuk 2 m.

Kata Kunci: Rekognisi Wajah, Oklusi Wajah, *Faster R-CNN*



ABSTRACT

NURINA RAHAYU. *Real Time Face Recognition System With Use Of Mask Occlusion* (supervised by Indrabayu and Elly Warni)

The use of masks today which is considered normal causes it to be difficult to recognise a person's face in public places so that this situation is used by suspects to camouflage so that their identities are not recognised, especially when still using conventional methods by matching suspect data with data that has been stored in the database so that it requires a long time and is inefficient.

The purpose of this research is to find out how the face recognition system with occlusion using masks in real time and how this system works both with occluded face and without occluded face.

In creating a face recognition system, the researcher uses Faster R-CNN (Region Based Convolutional Neural Networks) as a detection and classification system. If the system detects a face, both with and without occlusion, then the output is given in the form of classification results according to the classified face. The system performance is analysed using Multi-Class Confusion Matrix with a matrix size of 6x6 according to the number of classes in the system. There are two types of tests, which are Test 1 with 160 images and Test 2 conducted in real time in a conditioned classroom.

Test 1 resulted in an accuracy value of 100%. Test 2 with two scenarios based on two different distances, which are 3 m and 2 m with seven facial occlusion conditions, namely no occlusion, black KF94 mask occlusion, white KF94 mask occlusion, black duckbill mask occlusion, white duckbill mask occlusion, black regular mask occlusion, and white regular mask occlusion resulted in consecutive accuracies of 97,64%, 95,56%, 87,18%, 99,13%, 95,37%, 94,51% and 97,31% for a distance of 3 m and 99,8%, 100%, 99,3%, 99,49%, 99,81%, 99,63% and 100% for 2 m.

Keywords: Face Recognition, Face Occlusion, Faster R-CNN



DAFTAR ISI

LEMBAR PENGESAHAN SKRIPSI.....	i
PERNYATAAN KEASLIAN.....	ii
ABSTRAK.....	iii
ABSTRACT.....	iv
DAFTAR ISI.....	v
DAFTAR GAMBAR.....	vii
DAFTAR TABEL.....	ix
DAFTAR LAMPIRAN.....	xi
KATA PENGANTAR.....	xii
BAB 1 PENDAHULUAN.....	14
1.1 Latar Belakang.....	14
1.2 Rumusan Masalah.....	15
1.3 Tujuan Penelitian.....	15
1.4 Manfaat Penelitian.....	15
1.5 Ruang Lingkup.....	16
BAB 2 TINJAUAN PUSTAKA.....	17
2.1 Daftar Pencarian Orang (DPO), Orang Dalam Pemantauan (ODP) dan Pasien Dalam Pengawasan (PDP).....	17
2.2 Visi Komputer.....	17
2.2.1 Augmentasi Gambar.....	18
2.2.2 <i>Bounding box</i>	19
2.2.3 <i>Anchor boxes</i>	19
2.3 <i>Face Detection</i>	20
2.3.1 <i>DNN Face Detector</i>	20
2.4 <i>Face Recognition</i>	20
2.5 Oklusi.....	22
2.6 <i>Region Based CNN</i>	23
2.6.1 R-CNN.....	23
2.6.2 <i>Fast R-CNN</i>	24
2.6.3 <i>Faster R-CNN</i>	26
2.7 <i>Residual Network</i>	29
2.8 <i>Confusion Matrix</i>	31
2.8.1 <i>Precision</i>	32
2.8.2 Recall.....	33
2.8.3 <i>F1-Score</i>	33
2.8.4 Akurasi.....	33
2.8.5 Menentukan Nilai pada Kelas A.....	35
2.8.6 Menentukan Nilai pada Kelas B.....	35
2.8.7 Menentukan Nilai pada Kelas C.....	36
BAB 3 METODE PENELITIAN/PERANCANGAN.....	37
2.1 Tahapan Penelitian.....	37
2.2 Jenis Penelitian.....	38
2.3 Instrumen Penelitian.....	38
2.4 Teknik Pengambilan Data.....	39
2.5 Perancangan Sistem.....	42



3.5.1 Data <i>Input</i>	43
3.5.2 Data <i>preparation</i>	44
3.5.3 Implementasi <i>Faster R-CNN</i>	48
3.6 Evaluasi Sistem	50
BAB 4 HASIL DAN PEMBAHASAN.....	54
4.1 Hasil Penelitian	54
4.1.1 Pengujian 1	54
4.1.2 Pengujian 2	55
4.2 Pembahasan.....	77
BAB 5 KESIMPULAN DAN SARAN.....	89
5.1 Kesimpulan	89
5.2 Saran.....	90
DAFTAR PUSTAKA	91
LAMPIRAN.....	95



DAFTAR GAMBAR

Gambar 1 Model R-CNN	23
Gambar 2 Model <i>Fast R-CNN</i>	24
Gambar 3 <i>2x2 Region of interest pooling layer</i>	26
Gambar 4 Model <i>Faster R-CNN</i>	27
Gambar 5 <i>Faster R-CNN basic operation process</i>	28
Gambar 6 Diagram prinsip RPN	28
Gambar 7 Arsitektur <i>ResNet50</i>	29
Gambar 8 (a) Blok reguler dan (b) blok <i>residual</i>	30
Gambar 9 <i>Multi-class Confusion Matrix</i>	32
Gambar 10 Menghitung Nilai TP, FN, FP, dan TN pada <i>Confusion Matrix</i>	32
Gambar 11 Nilai <i>Confusion Matrix</i> pada kelas A.....	35
Gambar 12 Nilai <i>Confusion Matrix</i> pada kelas B	35
Gambar 13 Nilai <i>Confusion Matrix</i> pada kelas C	36
Gambar 14 Tahapan penelitian	37
Gambar 15 Ilustrasi pengambilan <i>data training</i>	39
Gambar 17 Ilustrasi pengambilan data uji	42
Gambar 18 Diagram blok sistem.....	43
Gambar 19 Anotasi pada gambar.....	45
Gambar 20 <i>Resize gambar</i> : (a) gambar <i>data training</i> 1920 x 1080 <i>pixel</i> dan (b) gambar <i>data training</i> setelah dilakukan <i>resize</i> menjadi 1280 x 720 <i>pixel</i>	45
Gambar 21 <i>Augmentasi flip</i> : (a) <i>data training</i> sebelum diaugmentasi dan (b) <i>data training</i> setelah diaugmentasi <i>flip</i>	46
Gambar 22 <i>Augmentasi rotasi gambar</i> berdasarkan garis horizontal: (a) <i>augmentasi rotasi gambar</i> pada -15° dan (b) <i>augmentasi rotasi</i> <i>gambar</i> pada 15°	46
Gambar 23 <i>Augmentasi saturation</i> : (a) contoh <i>data training</i> sebelum <i>augmentasi saturation</i> , (b) contoh <i>data training</i> setelah <i>augmentasi saturation -25%</i> dan (c) contoh <i>data training</i> setelah <i>augmentasi saturation 25%</i>	46
Gambar 24 <i>Augmentasi exposure</i> : (a) contoh <i>data training</i> sebelum <i>augmentasi exposure</i> , (b) contoh <i>data training</i> setelah <i>augmentasi</i> <i>augmentasi exposure -20%</i> dan (c) contoh <i>data training</i> setelah <i>augmentasi exposure 20%</i>	47
Gambar 25 <i>Augmentasi grayscale</i>	47
Gambar 26 <i>Augmentasi brightness</i> : (a) <i>Gambar data training</i> sebelum diaugmentasi dan (b) <i>gambar data training</i> setelah menggunakan <i>augmentasi brightness 15%</i>	48
Gambar 27 Arsitektur <i>Faster R-CNN</i>	49
Gambar 28 (a) Posisi orang pada pengujian 2 skenario 1 dan (b) Posisi orang pada pengujian 2 skenario 2	52
29 Matriks 6x6 <i>Multi-Class Confusion Matrix</i>	52
30 Perbandingan nilai <i>precision</i> , <i>recall</i> , dan <i>F1-score</i> pada pengujian 1	78



Gambar 31 Perbandingan nilai rata-rata *F1-score* untuk setiap kelas pada
pengujian 2 85



DAFTAR TABEL

Tabel 1 Tinggi badan subjek penelitian	40
Tabel 2 Contoh kondisi wajah pada <i>dataset</i> gambar	40
Tabel 3 Parameter <i>training</i> model <i>Faster R-CNN</i>	48
Tabel 4 Evaluasi pada penelitian.....	51
Tabel 5 Hasil <i>Confusion Matrix</i> pada pengujian 1.....	54
Tabel 6 Hasil pengujian 1	55
Tabel 7 Hasil <i>Confusion Matrix</i> pada pengujian 2 skenario 1 dengan oklusi tanpa masker	55
Tabel 8 Hasil pengujian 2 skenario 1 dengan oklusi tanpa masker	56
Tabel 9 Hasil <i>Confusion Matrix</i> pada pengujian dua skenario satu pada oklusi masker KF94 berwarna hitam.....	57
Tabel 10 Hasil pengujian 2 skenario 1 dengan oklusi masker KF94 berwarna hitam	57
Tabel 11 Hasil <i>Confusion Matrix</i> pada pengujian 2 skenario 1 dengan oklusi masker KF94 putih	59
Tabel 12 Hasil pengujian 2 skenario 1 dengan oklusi masker KF94 berwarna putih.....	59
Tabel 13 Hasil <i>Confusion Matrix</i> pada pengujian 2 skenario 1 pada oklusi masker duckbill berwarna hitam.....	60
Tabel 14 Hasil pengujian 2 skenario 1 dengan oklusi masker duckbill berwarna hitam	60
Tabel 15 Hasil <i>Confusion Matrix</i> pengujian 2 skenario 1 pada masker duckbill berwarna putih.....	62
Tabel 16 Hasil pengujian 2 skenario 1 dengan oklusi masker duckbill berwarna putih.....	62
Tabel 17 Hasil <i>Confusion Matrix</i> pengujian 2 skenario 1 pada oklusi masker reguler berwarna hitam.....	63
Tabel 18 Hasil pengujian 2 skenario 1 dengan oklusi masker reguler berwarna hitam	64
Tabel 19 Hasil <i>Confusion Matrix</i> pengujian 2 skenario 1 pada oklusi masker reguler putih.....	65
Tabel 20 Hasil pengujian 2 skenario 1 dengan oklusi masker reguler putih	65
Tabel 21 Hasil <i>Confusion Matrix</i> pada pengujian 2 skenario 2 dengan oklusi tanpa masker	67
Tabel 22 Hasil pengujian 2 skenario 2 dengan oklusi tanpa masker	67
Tabel 23 Hasil <i>Confusion Matrix</i> pada pengujian 2 skenario 2 pada oklusi masker KF94 berwarna hitam.....	68
Tabel 24 Hasil pengujian 2 skenario 2 dengan oklusi masker KF94 berwarna hitam	68
Tabel 25 Hasil <i>Confusion Matrix</i> pada pengujian 2 skenario 2 dengan oklusi masker KF94 berwarna putih	70
Hasil pengujian 2 skenario 2 dengan oklusi masker KF94 berwarna putih.....	70
Hasil <i>Confusion Matrix</i> pada pengujian 2 skenario 2 pada oklusi masker duckbill berwarna hitam.....	71



Tabel 28 Hasil pengujian 2 skenario 2 dengan oklusi masker duckbill berwarna hitam	72
Tabel 29. Hasil <i>Confusion Matrix</i> pengujian 2 skenario 2 pada masker duckbill berwarna putih	73
Tabel 30 Hasil pengujian 2 skenario 2 dengan oklusi masker duckbill berwarna putih	73
Tabel 31 Hasil <i>Confusion Matrix</i> pengujian 2 skenario 2 pada oklusi masker reguler berwarna hitam	74
Tabel 32 Hasil pengujian 2 skenario 2 dengan oklusi masker reguler berwarna hitam	75
Tabel 33 Hasil <i>Confusion Matrix</i> pengujian 2 skenario 2 pada oklusi masker reguler putih.....	76
Tabel 34 Hasil pengujian 2 skenario 2 dengan oklusi masker reguler putih	76
Tabel 35 Contoh Kesalahan pengenalan wajah pada pengujian 2	86



DAFTAR LAMPIRAN

Lampiran 1 Contoh Dataset Pengujian 1	95
Lampiran 2 Contoh Dataset Pengujian 2	96
Lampiran 3 <i>Source Code</i>	97



KATA PENGANTAR

Segala puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa, yang telah memberikan rahmat dan karunia-Nya, sehingga dapat menyelesaikan tugas akhir dengan judul “Sistem Rekognisi Wajah dengan Oklusi Penggunaan Masker Secara *Real Time*” sebagai salah satu syarat dalam menyelesaikan jenjang Strata-1 di Departemen Teknik Informatika, Fakultas Teknik, Universitas Hasanuddin.

Penulis menyadari bahwa penyusunan dan penulisan tugas akhir ini tidak dapat terselesaikan dengan baik tanpa adanya bantuan, bimbingan, serta dukungan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan ucapan terima kasih banyak kepada:

1. Kedua orang tua penulis, Bapak Alm. Aris Wahyudi dan Ibu Hamisah Hidjaz yang selalu memberikan dukungan, doa, dan semangat yang tiada hentinya, serta selalu sabar dalam mendidik penulis sejak kecil.
2. Bapak Prof. Dr. Ir. Indrabayu, ST., MT., M.Bus.Sys., IPM, ASEAN. Eng., selaku pembimbing I dan Ibu Elly Warni, S.T., M.T., selaku pembimbing II, yang senantiasa menyediakan waktu, tenaga, pikiran, dan perhatian yang luar biasa dalam mengarahkan penulis untuk menyelesaikan tugas akhir.
3. Bapak Prof. Dr. Ir. Indrabayu, ST., MT., M.Bus.Sys., IPM, ASEAN. Eng., selaku ketua Departemen Teknik Informatika Fakultas Teknik Universitas Hasanuddin.
4. Segenap keluarga *AIMP Research Group* Universitas Hasanuddin yang telah memberikan begitu banyak bantuan selama penelitian, pengambilan data dan diskusi progres penyusunan tugas akhir serta selalu memberikan motivasi.
5. Muhammad Fadhil, Muh. Ikhwan Ramadhani, Herlina Anwar, Irma Jufri, Devy, Muhammad Yusuf, Saphira Noer Sakinah dan Muh. Wahyudi R amara serta seluruh anggota Recognizer yang telah membantu penulis sejak awal perkuliahan dan selalu membantu dalam penyelesaian tugas akhir.



6. Segenap staf dan dosen Departemen Teknik Informatika, Fakultas Teknik, Universitas Hasanuddin yang telah membantu kelancaran penyelesaian tugas akhir.
7. Serta pihak-pihak lain yang tidak disebutkan dan tanpa sadar telah menjadi inspirasi dan membantu penulis dalam menyelesaikan tugas akhir.

Penulis berharap semoga Tuhan Yang Maha Esa berkenan membalas segala kebaikan serta jasa dari semua pihak yang telah banyak membantu penulis dalam menyelesaikan tugas akhir ini. Penulis menyadari bahwa tugas akhir ini masih jauh dari kata sempurna, oleh karena itu penulis mengharapkan segala bentuk saran serta masukan yang membangun dari berbagai pihak. Semoga tugas akhir ini dapat memberi manfaat bagi para pembaca dan semua pihak. Aamiin.

Makassar, 15 Maret 2024

Penulis



BAB 1

PENDAHULUAN

1.1 Latar Belakang

Sejak wabah COVID-19 merebak, penggunaan masker telah menjadi kebiasaan yang diterapkan secara luas dalam kehidupan sehari-hari. Kini, masker diwajibkan ketika keluar rumah atau berada di tempat umum guna mengurangi risiko penyebaran virus SARS-CoV-2. Berbagai organisasi kesehatan, termasuk Organisasi Kesehatan Dunia (WHO), menganjurkan penerapan aturan penggunaan masker, terutama di wilayah-wilayah di mana transmisi virus secara komunitas masih berlangsung. Berdasarkan WHO dan memperhatikan perkembangan penyakit di Indonesia, Kementerian Kesehatan Republik Indonesia bersama Gugus Tugas Percepatan Penanganan COVID-19 mengimbau kepada setiap pimpinan daerah untuk membuat kebijakan yang mewajibkan penggunaan masker oleh semua orang ketika berada di luar rumah dan himbauan penyediaan sarana cuci tangan pakai sabun (Kementerian Kesehatan Republik Indonesia, 2022).

Pada Rabu, 21 Juni 2023, pemerintah resmi mencabut status pandemic Covid-19 di Indonesia dan beralih menjadi endemik, sesuai dengan pengumuman yang disampaikan melalui video yang diunggah Sekretariat Presiden di Youtube (Sekretariat Presiden, 2023). Walaupun penggunaan masker tidak diwajibkan lagi, namun banyak masyarakat yang tetap menggunakan masker terutama di tempat umum sehingga penggunaan masker saat ini sudah menjadi kebiasaan yang melekat dalam kehidupan sehari-hari.

Karena penggunaan masker saat ini yang dianggap wajar menyebabkan sulitnya mengenali wajah seseorang di tempat umum sehingga keadaan ini dimanfaatkan oleh suspek, seperti yang termasuk dalam daftar pencarian orang (DPO), orang dalam pemantauan (ODP), maupun pasien dalam pengawasan (PDP) untuk berkamufase agar tidak dikenali identitasnya. Apalagi dalam

suspek masih menggunakan cara konvensional dengan mencocokkan data dengan data yang telah disimpan di *database* sehingga membutuhkan waktu yang cukup lama dan tidak efisien.



Penelitian yang dilakukan Huang menganalisis sistem rekognisi wajah dengan penggunaan masker menggunakan *deep learning multi-granularity recognition model* dengan menggunakan tiga jenis dataset gambar wajah pengguna masker menghasilkan akurasi pengenalan mencapai 95%, tetapi belum mencapai hasil yang maksimal dibandingkan teknologi pengenalan wajah biasa yang akurasinya lebih dari 99% (Huang dkk., 2021). Halawa membangun sistem rekognisi wajah dengan metode *Faster R-CNN* untuk mengidentifikasi pemilik kendaraan bermotor di halaman parkir menggunakan CCTV dengan akurasi 72,5% (Halawa dkk., 2019). Ding membangun sistem deteksi penggunaan masker menggunakan algoritma *Faster R-CNN ResNet50* yang hasilnya meningkatkan sistem deteksi terhadap penggunaan masker yang tidak benar (Ding dkk., 2022). Oleh karena itu, penulis mengusulkan judul “Sistem Rekognisi Wajah dengan Oklusi Penggunaan Masker Secara Real Time”. Sistem ini dirancang untuk mengenali wajah dengan menggunakan masker maupun tanpa menggunakan masker secara *real time* menggunakan metode *Faster R-CNN*.

1.2 Rumusan Masalah

Adapun rumusan masalah dari penelitian ini adalah:

1. Bagaimana merancang sistem pengenalan wajah secara *real time* dengan oklusi penggunaan masker?
2. Bagaimana unjuk kerja sistem rekognisi dengan oklusi maupun tanpa oklusi?

1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah:

1. Sistem melakukan rekognisi wajah dengan oklusi penggunaan masker secara *real time*.
2. Mengetahui performa sistem rekognisi dengan oklusi maupun tanpa oklusi.



Manfaat Penelitian

Manfaat yang dapat diperoleh dari penelitian ini adalah:

1. Memberikan informasi mengenai performa sistem rekognisi dengan oklusi maupun tanpa oklusi.
2. Membantu pengembangan rekognisi wajah dengan oklusi masker.
3. Memberikan informasi terhadap bidang ilmu pengetahuan yang relevan sesuai tema.

1.5 Ruang Lingkup

Ruang Lingkup dalam penelitian ini adalah:

1. Oklusi hanya berupa masker wajah saja, tidak ada oklusi tambahan.
2. Data diambil pada siang hari.
3. Hanya oklusi masker wajah yang tidak bergambar fitur wajah.



BAB 2 TINJAUAN PUSTAKA

2.1 Daftar Pencarian Orang (DPO), Orang Dalam Pemantauan (ODP) dan Pasien Dalam Pengawasan (PDP)

Menurut KBBI, Daftar Pencarian Orang merupakan orang yang sedang dicari keberadaannya karena melarikan diri dari kasus yang menyimpannya. Orang Dalam Pemantauan (ODP) merupakan seseorang yang mengalami gejala demam ($>38^{\circ}\text{C}$) atau ada riwayat demam atau ISPA tanpa pneumonia dan memiliki riwayat perjalanan ke negara yang terjangkit pada empat belas hari terakhir sebelum timbul gejala. Pasien dalam Pengawasan (PDP) merupakan seseorang yang memiliki riwayat kontak dengan kasus konfirmasi positif COVID-19, memiliki riwayat perjalanan ke negara yang terjangkit pada empat belas hari terakhir sebelum timbul gejala, dan memiliki riwayat demam dan satu dari gejala berikut: batuk, pilek, atau sesak napas tanpa disertai pneumonia (Kementerian Kesehatan RI, 2020).

Seseorang dapat termasuk dalam status Daftar Pencarian Orang (DPO) jika orang yang sudah ditetapkan sebagai tersangka oleh kepolisian melarikan diri (Ananta, 2019). Menurut Pasal 1 angka 14 KUHP menyatakan bahwa tersangka adalah seseorang yang karena perbuatannya atau keadaannya berdasarkan bukti permulaan patut diduga sebagai pelaku tindak pidana. Berdasarkan Peraturan Kepala Kepolisian Negara Republik Indonesia Nomor 14 Tahun 2012 tentang Manajemen Penyidikan Tindak Pidana dalam Pasal 31 ayat (1) bahwa tersangka yang telah dipanggil untuk pemeriksaan dalam rangka penyidikan perkara sampai lebih dari tiga kali dan ternyata tidak jelas keberadaannya, dapat dicatat di dalam Daftar Pencarian Orang (DPO) dan dibuatkan Surat Pencarian Orang.

2.2 Visi Komputer

Visi Komputer (*Computer Vision*) merupakan salah satu dari *artificial ce* yang memungkinkan komputer untuk melihat dan mengenali objek disekitarnya layaknya cara kerja mata manusia. Penerapan visi komputer emukan pada kehidupan kita dan saling berkaitan erat, seperti pada



medical diagnosis, self-driving vehicles, camera monitoring, atau smart filters. Dalam beberapa tahun terakhir, pembelajaran mendalam telah menjadi kekuatan transformatif untuk memajukan kinerja sistem visi komputer sehingga dapat dikatakan bahwa aplikasi *computer vision* hampir tidak dapat dipisahkan dari *deep learning* (Zhang dkk., 2023).

2.2.1 Augmentasi Gambar

Augmentasi gambar menghasilkan beberapa contoh yang serupa tetapi berbeda dengan melakukan perubahan secara acak pada *training set*. Dengan kata lain, augmentasi gambar digunakan dengan melakukan perubahan acak dari contoh pelatihan guna memungkinkan model dapat meningkatkan kemampuan generalisasinya pada objek sehingga model tidak bergantung pada atribut objek tertentu. Misalnya, sebuah citra gambar dipotong dengan berbagai bentuk untuk memunculkan objek yang diinginkan dengan posisi yang berbeda, sehingga model tidak bergantung pada posisi objek di citra gambar. Augmentasi juga dapat disesuaikan dari faktor-faktor citra tersebut seperti kecerahan dan warna untuk mengurangi sensitivitas model terhadap warna pada citra tersebut (Zhang dkk., 2023).

Saturation pada citra merupakan parameter yang mengukur sejauh mana warna dalam gambar terlihat jenuh atau kaya. Semakin tinggi nilai *saturation*-nya, warna akan terlihat lebih cerah dan hidup. Sementara semakin rendah nilainya, warna akan terlihat lebih pucat atau monokromatik (Taylor, 2021). Dalam augmentasi gambar, perubahan *saturation* dapat membantu model pembelajaran mesin dalam memahami variasi dan kondisi pencahayaan atau kualitas visual yang mungkin dapat dijumpai dalam aplikasi dunia nyata (Funabiki dkk., 2011).

Exposure pada citra merupakan parameter yang mengontrol seberapa terang atau gelap gambar tersebut. Hal ini berdasarkan banyak cahaya yang ditangkap oleh lensa kamera selama pengambilan gambar. Semakin tinggi nilai *exposure*-nya, maka akan membuat gambar menjadi lebih terang. Sedangkan semakin ilainya, maka akan membuat gambar lebih gelap. Pengaturan *exposure* ruh pada detail dan kontras dalam gambar (Taylor, 2021). Mengubah



exposure dapat membantu melatih model pembelajaran mesin untuk mengenali objek atau fitur dalam berbagai kondisi pencahayaan (Shin & Boyapati, 2022).

Flip dalam augmentasi gambar merujuk pada membalikkan atau diputar 180° secara horizontal atau vertikal. Tujuannya untuk menciptakan variasi tambahan dari data yang ada sehingga membantu meningkatkan kemampuan generalisasi dan *robustness* model *deep learning*, terutama dalam tugas pengenalan pola atau klasifikasi (Almattar & Algherairy, 2022).

Rotate pada augmentasi gambar merupakan teknik memutar gambar pada sudut tertentu. Hal ini dapat membantu model pembelajaran mesin untuk mengenali objek dari berbagai orientasi, meningkatkan ketahanan dan akurasi model pada kondisi dunia nyata (Xu dkk., 2020).

Grayscale pada augmentasi gambar merujuk pada proses mengubah gambar berwarna menjadi gambar skala abu-abu. Dalam konteks pembelajaran mesin dan pengolahan gambar, konversi ke *grayscale* sering digunakan sebagai teknik augmentasi data untuk meningkatkan variasi dalam set data pelatihan dan membantu model dalam memahami fitur penting gambar tanpa mengandalkan informasi warna (Longpre & Sohmsheety, 2016).

Brightness dalam konteks augmentasi gambar adalah salah satu teknik yang digunakan untuk mengubah tingkat kecerahan atau intensitas cahaya dalam gambar. Tujuannya adalah untuk meningkatkan variasi *data training*, membantu model jaringan saraf mendeteksi pola yang lebih umum, serta mencegah *overfitting*. Peningkatan cahaya dapat mencakup peningkatan atau penurunan tingkat kecerahan keseluruhan gambar (Goodfellow dkk., 2016).

2.2.2 *Bounding box*

Dalam pendeteksi objek, *bounding box* digunakan untuk menggambarkan lokasi spasial suatu objek pada citra. *Bounding box* berbentuk persegi panjang yang ditentukan oleh koordinat x dan y sebagai penanda batasan objek dari sisi atas, sisi kiri, sisi bawah dan sisi kanan objek (Zhang dkk., 2023).



Anchor boxes

Algoritma pendeteksi objek biasanya mengambil sebagian besar jumlah *region* dari gambar masukan untuk menentukan jika wilayah tersebut

berisi objek yang dikehendak, serta menyesuaikan batas wilayah untuk memprediksi objek dari *bounding box* secara benar dengan lebih akurat. Model yang berbeda memungkinkan penggunaan skema pengambilan sampel *region* yang berbeda. Salah satu metode tersebut mampu menghasilkan beberapa *bounding box* dari berbagai skala dan rasio aspek yang berpusat pada setiap piksel citra. Beberapa *bounding box* ini yang disebut *anchor boxes* (Zhang dkk., 2023).

2.3 Face Detection

Face Detection atau deteksi wajah adalah tugas penting dalam visi komputer dalam berbagai bidang seperti interaksi manusia-komputer, pengawasan, dan sistem biometric. *Face detection* adalah metode untuk mendeteksi wajah dari sebuah citra gambar yang memiliki beberapa fitur pada gambar tersebut. Hal ini melibatkan identifikasi dan lokalisasi wajah dalam gambar yang dapat menjadi tantangan karena variasi ukuran, bentuk, warna dan pencahayaan. Deteksi wajah menjadi tugas yang menantang karena karakteristik wajah yang tidak kaku dan dapat berubah dalam ukuran, bentuk, warna, dll. Deteksi wajah menjadi tugas yang lebih menantang ketika gambar yang diberikan tidak jelas, fitur wajah terhalang oleh benda lain, faktor pencahayaan yang tidak tepat dan wajah yang tidak menghadap kamera (Meena & Sharan, 2016).

2.3.1 DNN Face Detector

DNN (*Deep Neural Network*) *face detector* merupakan salah satu metode *face detection* berupa model Caffe yang dirancang dengan *Single Shot-Multibox Detector* (SSD) dan arsitekturnya adalah ResNet-10. Model Caffe ini diluncurkan dalam modul *deep neural network* (DNN) pada OpenCV 3.3. Model ini sangat andal, bekerja pada CPU secara *real time*, bekerja dengan baik pada berbagai orientasi wajah, bahkan pada oklusi yang signifikan dan mampu mendeteksi wajah dengan ukuran yang berbeda (Paul Jacob, 2021)



2.3.1.1 Face Recognition

Face Recognition atau pengenalan wajah merupakan salah satu varian *biometric* yang menerapkan sistem autentifikasi yang mana

pengembangannya sangat penting dalam beberapa dekade terakhir. *Face recognition* merupakan bagian dari pengenalan pola, dimana wajah seseorang dikategorikan sebagai dikenal atau tidak dikenal oleh sistem setelah membandingkan dengan gambar seseorang yang dikenali dari *database*. Pengenalan wajah merupakan sebuah tantangan yang memberikan variasi data tertentu karena beragam variasi data dari orang yang berbeda, termasuk variasi sistematis dari beberapa faktor seperti kondisi kecerahan dan pose seseorang (Hazim Barnouti dkk., 2016).

Metode komputasi untuk pengenalan wajah perlu mengatasi berbagai tantangan. Tantangan tersebut muncul karena wajah perlu ditampilkan dalam kondisi yang memungkinkan tampilnya informasi wajah untuk mengenali wajah secara spesifik dari seluruh data wajah yang ada di dalam *database*. Tampilan wajah merupakan masalah yang secara spesifik cukup sulit mengingat seluruh tampilan wajah hampir sama, terutama seluruh wajah memiliki dua mata, mulut, hidung, dan fitur wajah lainnya yang berada di posisi yang sama (Hazim Barnouti dkk., 2016).

Wajah manusia memiliki struktur yang dinamik dan kompleks dengan karakteristik yang bisa berubah secara signifikan seiring waktu. Pengenalan wajah melibatkan serangkaian aktivitas dalam berbagai aspek kehidupan manusia. Manusia bisa mengenali wajah, tetapi terkadang beberapa wajah sulit untuk diingat, pembelajaran mesin sekarang ditingkatkan untuk melakukan tugas ini. Para peneliti mencoba untuk memahami arsitektur wajah manusia ketika membangun atau mengembangkan sistem rekognisi wajah (Hazim Barnouti dkk., 2016).

Ketertarikan dan fokus dari metodologi sistem rekognisi wajah manusia dapat membantu para peneliti untuk memahami sistem dasarnya. Sistem rekognisi wajah manusia memanfaatkan beberapa data yang setidaknya berasal dari beberapa indra, seperti penglihatan, pendengaran, maupun sentuhan. Setiap data tersebut digunakan baik secara individual maupun bersamaan untuk mengingat

simpan data wajah. Dalam beberapa kasus, kondisi di sekitar orang juga hal yang penting dalam sistem rekognisi wajah manusia. Menangani data yang besar dan menggabungkannya merupakan hal yang sulit untuk



sebuah sistem rekognisi. Bagaimanapun, mengingat beberapa wajah juga hal yang sulit. Kunci utama dari sebuah sistem pembelajaran adalah kapasitas memori. Fitur wajah yang mungkin bisa digunakan untuk rekognisi wajah masih terus dipelajari. Dengan kata lain, diperlukan fitur umum dan spesifik untuk rekognisi wajah (Hazim Barnouti dkk., 2016).

2.5 Oklusi

Oklusi mengacu pada obstruksi atau hambatan satu objek atau bagian dari suatu objek dengan objek yang lainnya dalam suatu kondisi. Hal ini merupakan tantangan umum dalam visi komputer dan *object tracking*, dimana visibilitas objek atau bagian tertentu dari objek terhalangi oleh objek lain atau *self-occlusion*. Oklusi bias terjadi di berbagai kondisi, seperti ketika objek saling tumpang, ketika sebagian objek tertutupi, atau ketika objek dilihat dari sudut yang berbeda. Menangani oklusi dalam *computer vision* merupakan hal yang penting untuk akurasi *object detection*, *tracking*, dan *face recognition*. Beberapa teknik dan algoritma telah dikembangkan untuk mengatasi hambatan oklusi dalam pengaplikasian *computer vision* (Chandel & Vatta, 2015).

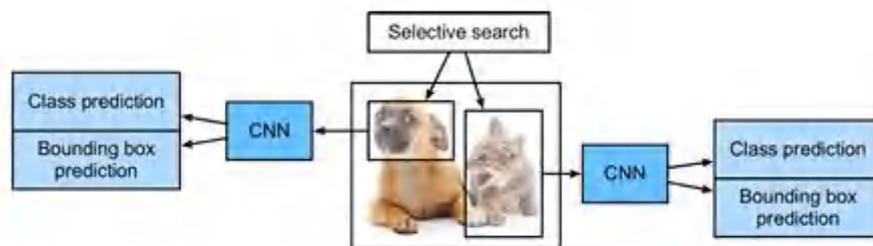
Dalam sistem pengenalan wajah, oklusi merujuk pada kondisi dimana bagian dari wajah seseorang tidak terlihat atau tertutupi dari pandangan. Hal ini disebabkan ketika objek atau bagian lain dari wajah menghalangi fitur wajah sehingga menjadi tantangan untuk mendeteksi secara akurat dan mengenali wajah seseorang. Oklusi bisa terjadi dalam beberapa kondisi, seperti ketika seseorang menggunakan kacamata, topi, masker, atau ketika seseorang menutupi wajah mereka dengan tangan atau objek lainnya. Menangani oklusi pada sistem pengenalan wajah merupakan tantangan yang spesifik dalam *computer vision* dan memerlukan pengembangan *robust algorithms* untuk mengidentifikasi dan mencocokkan fitur wajah secara akurat walaupun dengan adanya oklusi (Chandel & Vatta, 2015).



2.6 Region Based CNN

2.6.1 R-CNN

R-CNN pertama-tama mengekstraksi banyak, contohnya 2000 *region proposals* dari gambar *input* (misalnya, *anchor boxes* juga dapat dianggap sebagai *region proposals*), memberi label pada kelas dan *bounding box* (Girshick dkk., 2014). Kemudian CNN digunakan untuk melakukan *forward propagation* pada setiap *region proposals* untuk mengekstrak fitur-fiturnya. Selanjutnya, fitur-fitur dari *region proposals* digunakan untuk memprediksi kelas dan *bounding box* pada *region proposals* ini (Zhang dkk., 2023). Tampilan model R-CNN ditunjukkan pada Gambar 1.



Gambar 1 Model R-CNN

Sumber: (Zhang dkk., 2023)

Berdasarkan ilustrasi pada Gambar 1, R-CNN terdiri dari 4 tahap berikut:

1. Melakukan *selective search* untuk mengekstraksi beberapa *region proposals* berkualitas tinggi pada gambar *input*. Wilayah yang diusulkan ini biasanya dipilih pada beberapa skala dengan bentuk dan ukuran yang berbeda. Setiap proposal wilayah akan diberi label dengan kelas dan *bounding box*.
2. Memilih CNN yang telah dilatih sebelumnya dan potong sebelum lapisan keluaran. Ubah ukuran setiap proposal wilayah ke ukuran *input* yang diperlukan oleh jaringan, dan keluarkan fitur yang diekstraksi untuk proposal wilayah melalui *forward propagation*.
3. Mengambil fitur yang diekstraksi dan kelas berlabel dari setiap proposal wilayah sebagai contoh. Latih beberapa mesin vektor pendukung untuk klasifikasikan objek, di mana setiap mesin vektor pendukung secara dual menentukan apakah contoh berisi kelas tertentu.

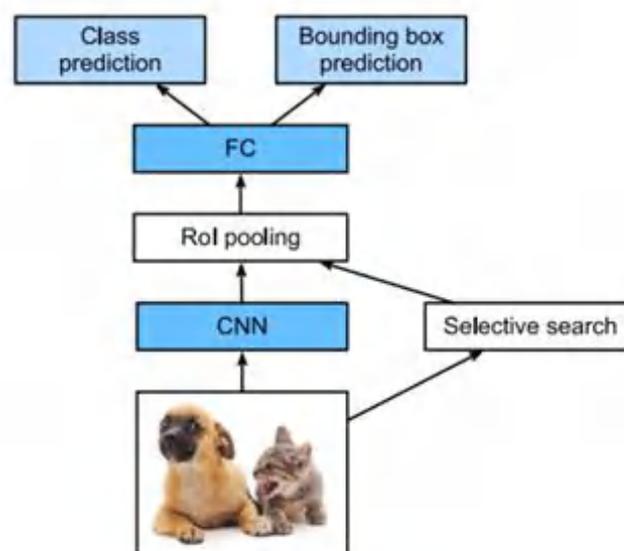


- Mengambil fitur yang diekstraksi dan *bounding box* berlabel setiap *region proposal* sebagai sampel. Latih model regresi linier untuk memprediksi *bounding box*.

Walaupun model R-CNN menggunakan CNN yang telah dilatih sebelumnya untuk mengekstraksi fitur gambar secara efektif, namun performanya lambat. Bayangkan kita memilih ribuan *region proposals* dari satu gambar *input*, dalam prosesnya dibutuhkan ribuan CNN *forward propagation* untuk melakukan deteksi objek. Karena membutuhkan beban komputasi yang besar membuat R-CNN tidak dapat digunakan secara luas dan *real world* (Zhang dkk., 2023).

2.6.2 Fast R-CNN

Hambatan performa R-CNN terletak pada *forward propagation* pada CNN untuk setiap *region proposal*, tanpa membagi komputasi. Karena *regions* ini biasanya saling tumpang tindih, ekstraksi fitur independen menyebabkan banyaknya komputasi yang berulang. Salah satu peningkatan besar *Fast R-CNN* dari R-CNN adalah bahwa CNN *forward propagation* dilakukan untuk setiap proposal objek pada seluruh citra dengan saling berbagi komputasi (Girshick dkk., 2014).



Gambar 2 Model *Fast R-CNN*



hang dkk., 2023)

Berdasarkan ilustrasi pada Gambar 2, komputasi utamanya adalah sebagai berikut:

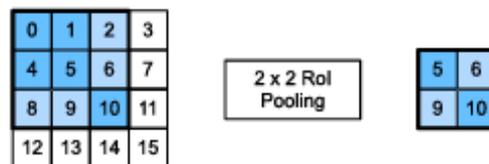
1. Dibandingkan dengan R-CNN, di dalam *Fast R-CNN input* dari CNN untuk ekstraksi fitur adalah seluruh citra, bukan *region proposals* individual. Apalagi CNN ini bisa dilatih. Diberikan sebuah *input* gambar, menghasilkan bentuk *output* CNN menjadi $1 \times c \times h1 \times w1$ dimana c , h dan w adalah *channel*, *height* dan *weight* dari *feature map*.
2. Misalkan *selective search* menghasilkan n *region proposals*. *Region proposals* ini (dalam bentuk yang berbeda) menandai *regions of interest* (dalam bentuk yang berbeda) pada *output* CNN. Lalu beberapa *regions of interest* ini mengekstrak fitur lebih lanjut dari bentuk yang sama (misalkan *height* $h2$ dan *weight* $w2$ ditentukan) agar dapat digabungkan dengan mudah. Untuk mencapai hal ini, *fast R-CNN* memperkenalkan *the region of interest (ROI) pooling layer*: *output* CNN dan *region proposals* dimasukkan pada lapisan ini, menghasilkan fitur gabungan dari bentuk $n \times c \times h2 \times w2$ yang diekstrak lebih lanjut untuk semua *region proposals*.
3. Menggunakan lapisan yang terhubung penuh, mentransformasikan gabungan fitur menjadi sebuah luaran dalam bentuk $n \times d$, dimana d bergantung pada desain model.
4. Memprediksi kelas dan *bounding box* untuk setiap n *region proposals*. Lebih jelasnya, pada masing-masing kelas dan prediksi *bounding box*, diubah luaran lapisan terhubung penuh menjadi sebuah luaran dalam bentuk $n \times q$ (dimana q adalah jumlah kelas) dan sebuah luaran dalam bentuk $n \times 4$. Prediksi kelas menggunakan *softmax regression*.

Pada *pooling layer*, kita tidak mengontrol secara langsung luaran bentuk dengan menentukan ukuran dari *pooling window*, *padding*, dan *stride*. Sebaliknya, kita bisa secara langsung menentukan luaran bentuk di *region of interest pooling layer* (Zhang dkk., 2023).



Sebagai contoh, kita tentukan masing-masing luaran tinggi dan lebar pada *region* sebagai $h2$ dan $w2$. untuk setiap daerah yang diinginkan jendela bentuk $h \times w$, jendela ini dibagi menjadi sebuah *grid subwindows* $h2 \times w2$, bentuk dari setiap *subwindows* diperkirakan $(h/h2) \times (w/w2)$. Dalam

praktiknya, tinggi dan lebar dari setiap *subwindows* harus dibulatkan ke atas, dan elemen yang paling besar harus digunakan sebagai luaran dari *subwindows*. Oleh karena itu, *region of interest pooling layer* bisa mengekstrak fitur dari bentuk yang sama walaupun ketika *region of interest* memiliki bentuk yang berbeda (Zhang dkk., 2023).



Gambar 3 2x2 *Region of interest pooling layer*

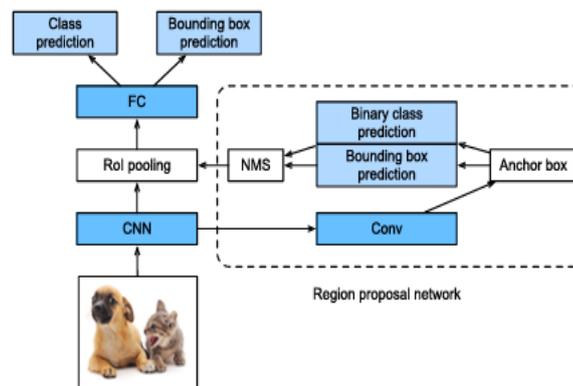
Sumber: (Zhang dkk., 2023)

Sebagai contoh, pada Gambar 3 bagian 3x3 kiri atas *region of interest* dipilih input 4x4. Untuk *region of interest* ini, digunakan 2x2 *region of interest pooling layer* untuk mendapatkan output 2x2. Ingat bahwa setiap empat bagian *subwindows* memiliki elemen 0, 1, 4 dan 5 (5 merupakan nilai maksimum); 2 dan 6 (6 merupakan nilai maksimum); 8 dan 9 (9 merupakan nilai maksimum); dan 10 (Zhang dkk., 2023).

2.6.3 *Faster R-CNN*

Agar lebih akurat dalam pendeteksian objek, model *Fast R-CNN* biasanya harus menghasilkan banyak *region proposals* pada *selective search*. Untuk mengurangi *region proposals* tanpa penurunan akurasi, *Faster R-CNN* dirancang untuk mengganti *selective search* dengan sebuah *region proposal network* (Zhang dkk., 2023).





Gambar 4 Model *Faster R-CNN*

Sumber: (Zhang dkk., 2023)

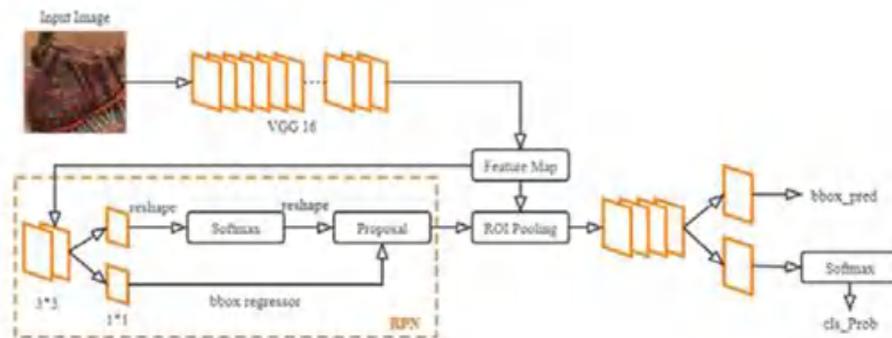
Dibandingkan dengan *Fast R-CNN*, *Faster R-CNN* hanya mengganti metode *region proposal* dari *selective search* menjadi sebuah *region proposal network*. Selain dari itu tidak ada perubahan pada model. Berdasarkan ilustrasi pada Gambar 4, *Region proposal network* bekerja dengan beberapa langkah berikut :

1. Menggunakan *convolutional layer 3x3* dengan *padding 1* untuk mengubah luaran CNN menjadi sebuah luaran baru dengan *c channels*. Dengan ini, setiap unit di sepanjang dimensi spasial *feature maps* yang diekstraksi mendapatkan sebuah vektor fitur baru dari panjang *c*
2. Berpusat pada setiap piksel dari *feature maps*, menghasilkan beberapa *anchor boxes* dari skala dan aspek rasio yang berbeda dan dilabeli.
3. Menggunakan panjang vektor fitur *c* pada pusat setiap *anchor box*, memprediksi kelas biner (latar belakang dan objek) dan *bounding box* untuk *anchor box* ini.
4. Pertimbangkan *bounding box* yang telah diprediksi yang kelas prediksinya adalah objek. Hilangkan hasil yang tumpang tindih menggunakan penekanan non-maksimum. Sisa prediksi *bounding box* untuk objek adalah *region proposals* yang diperlukan oleh *region of interest pooling layer*.

Perlu dicatat bahwa sebagai bagian dari model *Faster R-CNN*, *region proposal network* dilatih bersamaan dengan model lainnya. Dengan kata lain, tujuan dari *Faster R-CNN* tidak hanya pada kelas dan prediksi *bounding box* deteksi objek, tetapi juga pada kelas biner dan prediksi *bounding box* untuk objek pada *region proposal network*. Hasil dari pelatihan *end-to-end*,



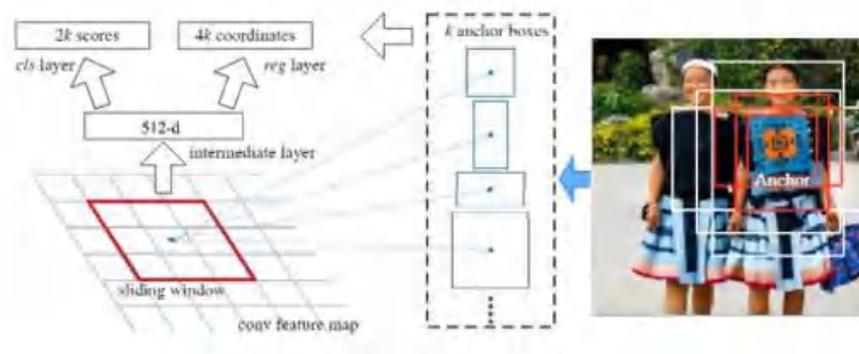
region proposal network mempelajari cara menghasilkan *region proposals* yang kualitas tinggi, sehingga tetap akurat dalam mendeteksi objek dengan mengurangi jumlah *region proposals* yang dipelajari dari data (Zhang dkk., 2023). Proses operasi dasar *Faster R-CNN* dapat dilihat pada Gambar 5.



Gambar 5 Faster R-CNN basic operation process

Sumber: (Kong dkk., 2020)

Faster R-CNN menyediakan metode konfigurasi VGG 16, Resnet_v1 dan *pre-trained models* lainnya. Pertama, *feature map* dihasilkan dari ekstraksi fitur pada kumpulan data gambar. Kemudian ditransfer ke *region proposal network* (RPN) untuk diproses. Proses utamanya yaitu: menghasilkan *anchors*, *softmax classification*, *bounding box regression*, dan proposal.



Gambar 6 Diagram prinsip RPN

Sumber: (Kong dkk., 2020)

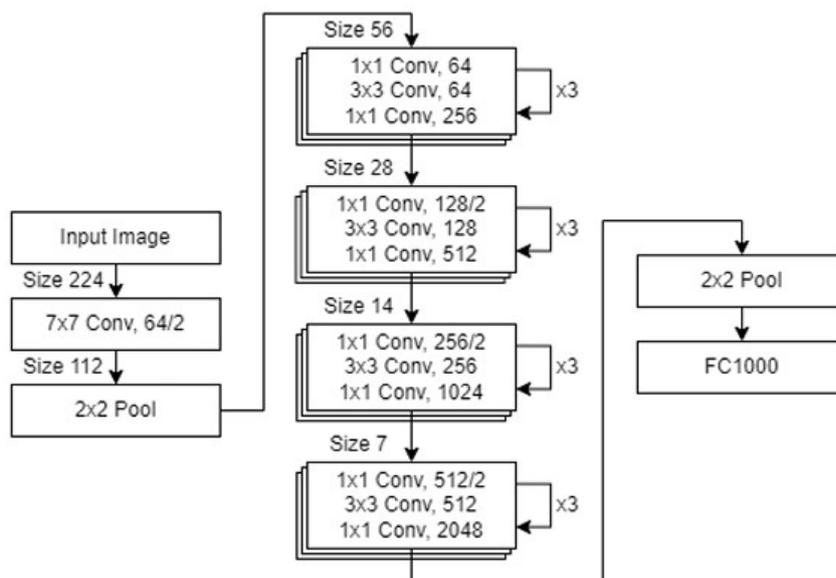
Berdasarkan Gambar 6, pada *feature map*, setiap piksel dapat dianggap sebagai sebuah *anchors*. RPN menghasilkan beberapa *anchor boxes* untuk setiap piksel dengan menyesuaikan koefisiennya. Kandidat area ini dikonversi ke *fully connected features* dengan konvolusi dan pemetaan dari fitur ke fitur. Klasifikasi dan lapisan regresi dihasilkan menggunakan *convolution*



kernels $2*k$ dan $4*k$. Pada lapisan klasifikasi, *softmax classification* digunakan untuk menentukan jika area kandidat termasuk ke latar depan (yaitu target). Pada lapisan regresi, regresi *bounding box* digunakan untuk modifikasi area kandidat, sehingga lebih dekat ke area yang sebenarnya (yaitu penandaan *box* selama *preprocessing*). Kemudian, lapisan proposal akan disesuaikan dengan poin *anchor* berdasarkan hasil regresinya, dan memilih beberapa area kandidat dengan nilai prospek yang tinggi sebagai proposal (Kong dkk., 2020).

Setelah itu, *ROI pooling layer* akan mengumpulkan proposal dan *feature maps* untuk menghasilkan *proposal feature maps*, Luaran dari *proposal feature maps* dihitung melalui *fully connected layer* dan klasifikasi *softmax* untuk menentukan kategori mana yang dimiliki setiap proposal, luaran vektor probabilitas *cls_prob*, dan mendapatkan posisi *offset bbox_pred* pada setiap proposal dengan regresi *bbox* untuk mengekstrak bingkai deteksi lebih akurat. Luaran parameter seperti *cls_prob* dan *bbox_pred* akan digunakan untuk mengoptimalkan model luaran latihan CNN dan meningkatkan performa deteksi *Faster R-CNN* (Kong dkk., 2020).

2.7 Residual Network



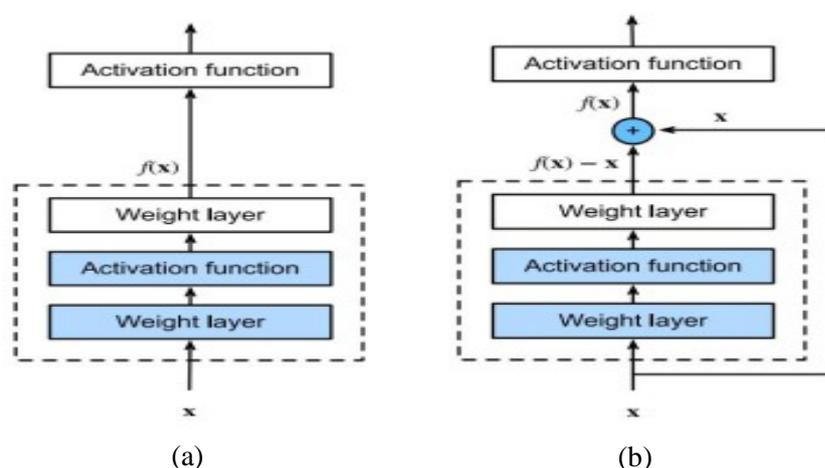
Gambar 7 Arsitektur *ResNet50*

er: (Bendjillali dkk., 2020)



Pada Gambar 7 menunjukkan lapisan konfigurasi jaringan pada ResNet50. Operasi yang paling penting pada CNN adalah *convolutional layer*, yang merupakan blok inti dari *Convolutional Neural Network* (CNN) yang melakukan sebagian besar komputasi berat. Salah satu contoh model *convolutional layer* yang cukup populer adalah ResNet. ResNet 50 adalah *residual network* yang terdiri dari 50 *layer*. Seperti VGG, ukuran *convolutional layers* adalah filter 3x3, ukuran masukan pada model ini ditetapkan sebagai 224x224 dan diikuti beberapa rancangan sederhana, seperti: Untuk lapisan yang memiliki jumlah filter yang sama memiliki luaran yang sama. Jumlah filter digandakan jika ukuran luaran yang terbelit dibelah dua sehingga kompleksitas waktu per lapisan dipertahankan. Model diakhiri dengan *pooling layer* rata-rata dan *1000-way-fully-connected layer* dengan *softmax*. Model ini memiliki lebih sedikit filter dan kompleksitas yang lebih rendah dibandingkan dengan VGG dan ada beberapa variasi lainnya seperti ResNet101 dan ResNet152 (Bendjillali dkk., 2020).

Inti dari *residual network* (ResNet) adalah gagasan bahwa setiap lapisan tambahan harus lebih mudah mengakomodasi fungsi identitas antar komponennya. Pertimbangan ini agak mendalam tetapi mengarah pada solusi yang sangat sederhana, yaitu *residual block* yang ditunjukkan pada Gambar 8.



Gambar 8 (a) Blok reguler dan (b) blok *residual*

hang dkk., 2023)



Berdasarkan pada Gambar 8, nilai *input* dinyatakan sebagai x . Bagan (a) pada Gambar 8, kotak yang terdapat di dalam garis putus-putus harus langsung mempelajari pemetaan $f(x)$. Bagan (b) pada Gambar 8, kotak yang terdapat di dalam garis putus-putus perlu mempelajari pemetaan *residual* $f(x) - x$, yang merupakan cara blok *residual* mendapatkan namanya. Jika pemetaan $f(x) = x$ adalah pemetaan dasar yang diinginkan, maka kita hanya perlu menentukan bobot dan bias lapisan bobot atas (misalnya, *fully connected layer* dan *convolutional layer*) pada kotak di dalam garis putus-putus menjadi nol. Bagan kanan pada Gambar 8 menggambarkan blok *residual* ResNet, dimana panah yang membawa *input* lapisan x ke operator tambahan disebut *sambungan residual* (atau *sambungan pintasan*). Dengan blok *residual*, *input* dapat menyebar lebih cepat melalui koneksi *residual* di seluruh lapisan (Zhang dkk., 2023).

ResNet mengikuti rancangan penuh lapisan konvolusi 3x3 VGG. Blok *residual* memiliki dua lapisan konvolusi 3x3 dengan jumlah saluran *output* yang sama. Setiap lapisan konvolusi diikuti oleh batch lapisan normalisasi dan fungsi aktivasi ReLU. Kemudian, lewati dua operasi konvolusi ini dan menambahkan *input* langsung sebelum fungsi aktivasi ReLU terakhir. Jenis desain ini mensyaratkan jika *output* dua lapisan konvolusi harus memiliki bentuk yang sama seperti *input*, sehingga bisa dijumlahkan. Jika ingin mengubah nilai saluran, maka perlu diperkenalkan lapisan konvolusi 1x1 tambahan untuk mengubah *input* menjadi bentuk yang diinginkan untuk operasi tambahan (Zhang dkk., 2023).

2.8 Confusion Matrix

Confusion Matrix adalah sebuah matriks yang menunjukkan jumlah kejadian antar dua nilai label, yaitu label aktual dan label prediksi. Sebuah *Confusion Matrix* memiliki tabel berukuran $l \times l$, dimana l adalah jumlah nilai label atau kelas yang berbeda (Kohavi & Provost, 1998). Pada *Multi-class Confusion Matrix*, kelas-kelas diurutkan dalam urutan yang sama dalam baris dan kolomnya seperti pada Gambar 9. Oleh karena itu, elemen yang diklasifikasi *ue Positive*) terletak pada diagonal utama dari kiri atas ke kanan bawah. Elemen yang merupakan titik temu antara prediksi dan aktual (Grandini dkk., 2020).



Berikut merupakan contoh *Multi-class Confusion Matrix* yang ditunjukkan pada Gambar 9.

		PREDICTED classification				Total
		Classes	a	b	c	
ACTUAL classification	a	6	0	1	2	9
	b	3	9	1	1	14
	c	1	0	10	2	13
	d	1	2	1	12	16
Total		11	11	13	17	52

Gambar 9 *Multi-class Confusion Matrix*

Sumber:(Grandini dkk., 2020)

2.8.1 Precision

		PREDICTED		Total
		Classes	Positive (1)	
ACTUAL	Positive (1)	TP = 20	FN = 5	25
	Negative (0)	FP = 10	TN = 15	25
Total		30	20	50

Gambar 10 Menghitung Nilai TP, FN, FP, dan TN pada *Confusion Matrix*

Sumber:(Grandini dkk., 2020)

Precision adalah nilai pecahan dari elemen *True Positive* dibagi dengan jumlah total unit yang diprediksi secara positif (jumlah kolom dari prediksi positif). Berdasarkan pada Gambar 10, persamaan *precision* dapat ditunjukkan pada persamaan (1) berikut,



$$recision = \frac{TP}{TP + FP} \times 100\% , \quad (1)$$

= *True Positive*,

FP = *False Negative*.

TP (*True Positive*) adalah elemen yang telah dilabeli sebagai positif pada model dan nilai sebenarnya positif, sedangkan FP (*False Positive*) adalah elemen yang telah dilabeli sebagai positif pada model, namun nilai sebenarnya negatif (Grandini dkk., 2020).

2.8.2 Recall

Recall adalah nilai pecahan dari elemen *True Positive* (TP) dibagi dengan jumlah total unit yang diklasifikasi positif (jumlah baris dari yang sebenarnya bernilai positif). Persamaannya dapat ditulis sebagai berikut,

$$Recall = \frac{TP}{TP + FN} \times 100\% , \quad (2)$$

dimana,

TP = *True Positive*,

FN = *False Negative*.

FN (*False Negative*) adalah elemen yang telah dilabeli sebagai negatif pada model, namun nilai sebenarnya positif (Grandini dkk., 2020).

2.8.3 F1-Score

F1-Score menilai kinerja model klasifikasi dari *Confusion Matrix*. Rumus *F1-Score* dapat diartikan sebagai rata-rata antara *precision* dengan *recall*, seperti persamaan berikut,

$$F1-Score = 2 \times \left(\frac{Precision \times Recall}{Precision + Recall} \right) \times 100\% , \quad (3)$$

di mana *F1-Score* mencapai nilai terbaiknya pada 1 dan skor terburuk pada 0. Kontribusi relative dari *precision* dan *recall* adalah sama pada *F1-Score* dan rata-rata harmonic berguna untuk menemukan *trade-off* terbaik antara dua kuantitas (Grandini dkk., 2020).

2.8.4 Akurasi



Akurasi merupakan salah satu matriks yang paling populer pada klasifikasi data. Akurasi menghitung pecahan jumlah dari elemen *True Positive* dan *True Negative* sebagai pembilang dan jumlah seluruh entri *Confusion*

Matrix yang ada pada Gambar 10 sebagai penyebut. Persamaannya dapat ditulis sebagai berikut,

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% , \quad (4)$$

dimana,

TP = *True Positive*,

TN = *True Negative*,

FP = *False Positive*,

FN = *False Negative*.

TP (*True Positive*) dan TN (*True Negative*) adalah elemen yang terklasifikasi benar pada model dan terletak pada diagonal utama *Confusion Matrix*, sedangkan yang sebagai penyebut adalah semua elemen termasuk yang diluar diagonal utama yang terklasifikasi salah oleh model, yaitu nilai TP, TN, FP, dan FN (Grandini dkk., 2020).

Akurasi mengukur seberapa benar model mampu memprediksi keseluruhan data secara benar. Elemen dasar matriksnya adalah individu tunggal dalam kumpulan data, yaitu setiap unit memiliki bobot yang sama dan berkontribusi setara pada nilai akurasi. Ketika berpikir tentang kelas daripada individu tunggal, akan ada kelas yang memiliki nilai unit yang lebih tinggi dan yang lainnya memiliki nilai unit yang sedikit. Dalam keadaan ini, kelas dengan populasi yang lebih tinggi memiliki bobot yang lebih tinggi dibandingkan kelas yang lebih rendah. Oleh karena itu, akurasi lebih cocok ketika kita mempertimbangkan individu tunggal daripada beberapa kelas (Grandini dkk., 2020).

Dalam menentukan nilai setiap sel pada *Confusion Matrix*, nilai TP, FP, FN dan TN berbeda untuk setiap kelas. Berikut cara dalam menentukan nilai pada *Confusion Matrix*:



2.8.5 Menentukan Nilai pada Kelas A

Tampilan nilai *Confusion Matrix* pada kelas A ditunjukkan pada Gambar 11 berikut.

		Predicted Classification		
		Classes	a	b
Artificial Classification	a	TP	FN	FN
	b	FP	TN	TN
	c	FP	TN	TN

Gambar 11 Nilai *Confusion Matrix* pada kelas A

Berdasarkan pada Gambar 11, kelas sel didefinisikan dengan kolom prediksi dan baris aktual. Nilai TP untuk kelas A terdapat pada sel (a, a). Nilai FP untuk kelas A yaitu seluruh sel pada kolom prediksi (a) yang bukan bernilai TP, yaitu sel (a, b) dan (a, c). Nilai FN untuk kelas A yaitu seluruh sel pada baris aktual (a) yang bukan bernilai TP, yaitu sel (b, a) dan (c, a). Nilai TN untuk kelas A yaitu seluruh sel tabel yang bukan bernilai TP, FP dan FN, yaitu sel (b, b), (c, b), (b, c), dan (c, c).

2.8.6 Menentukan Nilai pada Kelas B

Tampilan nilai *Confusion Matrix* pada kelas B ditunjukkan pada Gambar 12 berikut.

		Predicted Classification		
		Classes	a	b
Artificial Classification	a	TN	FP	TN
	b	FN	TP	FN
	c	TN	FP	TN

Gambar 12 Nilai *Confusion Matrix* pada kelas B

dasarkan pada Gambar 12, kelas sel didefinisikan dengan kolom prediksi aktual. Nilai TP untuk kelas B terdapat pada sel (b, b). Nilai FP untuk



kelas B yaitu seluruh sel pada kolom prediksi (b) yang bukan bernilai TP, yaitu sel (b, a) dan (b, c). Nilai FN untuk untuk kelas B yaitu seluruh sel pada baris aktual (b) yang bukan bernilai TP, yaitu sel (a, b) dan (c, b). Nilai TN untuk kelas B yaitu seluruh sel tabel yang bukan bernilai TP, FP dan FN, yaitu sel (a, a), (c, a), (a, c), dan (c, c).

2.8.7 Menentukan Nilai pada Kelas C

Tampilan nilai *Confusion Matrix* pada kelas C ditunjukkan pada Gambar 13 berikut.

		Predicted Classification		
		Classes	a	b
Artificial Classification	a	TN	TN	FP
	b	TN	TN	FP
	c	FN	FN	TP

Gambar 13 Nilai *Confusion Matrix* pada kelas C

Berdasarkan pada Gambar 13, kelas sel didefinisikan dengan kolom prediksi dan baris aktual. Nilai TP untuk kelas C terdapat pada sel (c, c). Nilai FP untuk kelas C yaitu seluruh sel pada kolom prediksi (c) yang bukan bernilai TP, yaitu sel (c, a) dan (c, b). Nilai FN untuk untuk kelas C yaitu seluruh sel pada baris aktual (c) yang bukan bernilai TP, yaitu sel (a, c) dan (b, c). Nilai TN untuk kelas A yaitu seluruh sel tabel yang bukan bernilai TP, FP dan FN, yaitu sel (a, a), (b, a), (a, b), dan (b, b).

