

TESIS

DETEKSI LOGO DENGAN MENGGUNAKAN METODE *YOU ONLY LOOK ONCE* (YOLO)

Logo Detection Using You Only Look Once (YOLO) Method

KIKI REZKIANI

D032181012



**PROGRAM STUDI MAGISTER TEKNIK ELEKTRO
DEPARTEMEN TEKNIK ELEKTRO
FAKULTAS TEKNIK
UNIVERSITAS HASANUDDIN
GOWA
2023**

PENGAJUAN TESIS

**DETEKSI LOGO DENGAN MENGGUNAKAN METODE *YOU ONLY LOOK
ONCE (YOLO)***

Tesis

Sebagai Salah Satu Syarat Untuk Mencapai Gelar Magister
Program Studi Teknik Elektro

Disusun dan diajukan oleh

KIKI REZKIANI

D032181012

Kepada

**FAKULTAS TEKNIK
UNIVERSITAS HASANUDDIN
GOWA
2023**

TESIS
DETEKSI LOGO DENGAN MENGGUNAKAN METODE
YOU ONLY LOOK ONCE (YOLO)

KIKI REZKIANI
D032 181 012

Telah dipertahankan di hadapan Panitia Ujian Tesis yang dibentuk dalam rangka penyelesaian studi pada Program Magister Teknik Elektro Fakultas Teknik Universitas Hasanuddin pada tanggal 19 Januari 2023 dan dinyatakan telah memenuhi syarat kelulusan

Menyetujui,

Pembimbing Utama



Dr. Ir. Ingrid Nurtanio, M.T.
NIP. 19610813 198811 2 001

Pembimbing Pendamping



Prof. Dr. Eng. Ir. Syafaruddin, S.T., M.Eng., IPU.
NIP. 19740530 199903 1 003

Dekan Fakultas Teknik
Universitas Hasanuddin



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

Ketua Program Studi
S2 Teknik Elektro



Dr. Eng. Ir. Wardi, S.T., M.Eng.
NIP. 19720828 199903 1 003

PERNYATAAN KEASLIAN TESIS DAN PELIMPAHAN HAK CIPTA

Yang bertanda tangan di bawah ini

Nama : Kiki Rezkiani
Nomor Mahasiswa : D032181012
Departemen : Teknik Elektro

Dengan ini menyatakan bahwa, tesis berjudul “Deteksi Logo Dengan Menggunakan Metode *You Only Look Once* (YOLO)” adalah benar karya saya dengan arahan dari komisi pembimbing (Dr. Ir. Ingrid Nurtanio, M.T. sebagai Pembimbing Utama dan Prof. Dr. Eng. Ir. Syafaruddin, S.T., M.Eng., IPU. sebagai Pembimbing Pendamping). Karya ilmiah ini belum diajukan dan tidak sedang diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka tesis ini. Sebagian dari isi tesis ini telah dipublikasikan di Jurnal/Prosiding (2022 2nd International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS)) sebagai artikel dengan judul “*Logo Detection Using You Only Look Once (YOLO) Method*”.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya berupa tesis ini kepada Universitas Hasanuddin.

Gowa, 19 Januari 2023

Yang Menyatakan,



Kiki Rezkiani

KATA PENGANTAR

Segala puji hanya milik Allah Subhanahu Wa Ta'ala yang telah memberikan rahmat, hidayah, taufik dan pertolongan-Nya dalam menyelesaikan tesis yang berjudul “**Deteksi Logo Dengan Menggunakan Metode *You Only Look Once (YOLO)***” sebagai salah satu persyaratan untuk memperoleh gelar Magister Teknik (M.T.) di Program Pascasarjana Departemen Teknik Elektro, Fakultas Teknik Universitas Hasanuddin. Shalawat serta salam semoga tetap tercurahkan kepada Nabi Muhammad Shallallahu ‘alaihi Wa Sallam, beserta keluarga dan para sahabatnya yang telah membimbing kita dari jalan kegelapan menuju jalan yang terang benderang.

Dalam kesempatan ini penulis dengan tulus menyampaikan terima kasih kepada:

1. Ibu Dr. Ir. Ingrid Nurtanio, M.T. selaku ketua pembimbing sekaligus sebagai bagian dari tim penilai tesis, dan Bapak Prof. Dr. Eng. Ir. Syafaruddin, S.T., M.Eng., IPU. selaku wakil pembimbing sekaligus sebagai bagian dari tim penilai tesis. Terima kasih atas motivasi dan bimbingan yang telah diberikan sehingga penyusunan tesis ini dapat terselesaikan. Melalui beliau berdua yang penuh kesabaran dan perhatian, telah memberikan arahan-arahan dan saran maupun kritik yang membangun pada aspek penulisan, metodologi, serta isi dari tesis ini secara keseluruhan.
2. Bapak Dr. Ir. Zahir Zainuddin, M.Sc., Bapak Dr. Eng. Muhammad Niswar, S.T., M.IT., dan Ibu Dr. A. Ejah Umraeni Salam, S.T., M.T. selaku komisi tim penguji. Terima kasih atas segala saran dan masukan, serta kerelaannya dalam meluangkan waktu mendiskusikan hal-hal yang menjadi keterbatasan selama penyusunan tesis ini.
3. Rektor Universitas Hasanuddin Bapak Prof. Dr. Ir. Jamaluddin Jompa, M.Sc., Dekan Fakultas Teknik Bapak Prof. Dr. Eng. Ir. Muhammad Irsan Ramli, S.T., M.T., beserta staff dan jajarannya, Ketua Departemen Teknik Elektro Ibu Dr. Eng. Ir. Dewiani, M.T., Ketua Program Studi S2 Teknik Elektro Bapak Dr. Eng. Ir. Wardi, S.T., M.Eng., dan seluruh dosen Teknik Elektro Universitas Hasanuddin yang telah memberikan kesempatan dan

bekal ilmu kepada penulis yang sangat berarti dalam menyelesaikan studi sampai pada terselesaikannya tesis ini. Kepada seluruh staff S2 Teknik Elektro Universitas Hasanuddin yang telah memberikan layanan terbaik pada aspek administrasi selama penulis menjalani studi di Departemen Teknik Elektro Universitas Hasanuddin.

4. Keluarga, para sahabat, teman seangkatan, maupun sesama mahasiswa Teknik Elektro Universitas Hasanuddin. Terima kasih atas segala do'a dan dukungannya selama ini.

Penulis berharap semoga penelitian ini dapat memberikan kontribusi pada masyarakat dan perkembangan ilmu pengetahuan dan teknologi.

Gowa, Januari 2023

Penulis

ABSTRAK

KIKI REZKIANI. *Deteksi Logo Dengan Menggunakan Metode You Only Look Once (YOLO)* (dibimbing oleh **Ingrid Nurtanio** dan **Syafaruddin**)

Jumlah dokumen yang semakin meningkat memerlukan klasifikasi dokumen otomatis karena membutuhkan waktu yang lebih lama jika dilakukan secara manual. Logo memungkinkan untuk menentukan sumber dokumen dengan cepat dan akurat. Logo menjadi bagian dari identitas yang melekat pada perusahaan, organisasi, lembaga maupun individu. Tata letak logo pada setiap dokumen berbeda-beda, serta kualitas gambar dari sebuah dokumen berbeda-beda. Penelitian ini bertujuan untuk mendeteksi logo pada sebuah dokumen berupa ijazah perguruan tinggi. Dalam penelitian ini, digunakan ijazah dari 15 perguruan tinggi. Pada sistem ini diimplementasikan *Darknet Framework* dengan model *You Only Look Once* versi 4 (YOLOv4). Hasil model akhir YOLOv4 menunjukkan sistem dapat mendeteksi dan mengenali logo dengan hasil *mean Average Precision* (mAP) adalah 93.73 %.

Kata kunci: logo, deteksi objek, *deep learning*, *Darknet*, YOLO

ABSTRACT

KIKI REZKIANI. *Logo Detection Using You Only Look Once (YOLO) Method* (supervised by **Ingrid Nurtanio** and **Syafaruddin**)

The growing amount of documents necessitates automatic document classification since manual classification takes longer. Logos allow for defining the source of a document rapidly and accurately. Logos are part of the identity attached to companies, organizations, institutions, and individuals. The placement of the logo on each document varies, as does the image quality of a document. This research aimed to identify the logo on a document in the form of a college diploma. This study utilized diplomas from fifteen universities. This system utilizes the You Only Look Once algorithm version 4 (YOLOv4) by implementing the Darknet Framework. The ultimate YOLOv4 model's results indicated that the system could detect and recognize logos with a mean Average Precision (mAP) of 93.73%.

Keywords: logo, object detection, deep learning, Darknet, YOLO

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
PENGAJUAN TESIS.....	ii
PERSETUJUAN TESIS	iii
PERNYATAAN KEASLIAN TESIS	iv
KATA PENGANTAR.....	v
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian.....	3
1.4 Manfaat Penelitian.....	3
1.5 Batasan Masalah.....	4
1.6 Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA.....	6
2.1 Landasan Teori	6
2.1.1 Logo	6
2.1.2 Citra.....	7
2.1.3 Pengolahan citra.....	8
2.1.4 <i>Computer vision</i>	10
2.1.5 Pengenalan pola	10
2.1.6 Segmentasi citra	11
2.1.7 <i>Thresholding</i>	11
2.1.8 Augmentasi	12
2.1.9 Anotasi gambar	20
2.1.10 YOLO.....	25

2.1.11	<i>Google colaboratory</i>	26
2.2	Penelitian Terkait.....	28
2.3	<i>State of the Art</i>	29
2.4	Kerangka Pikir.....	31
BAB III METODOLOGI PENELITIAN		32
3.1	Tahapan Penelitian	32
3.2	Waktu dan Lokasi Penelitian.....	33
3.3	Instrumen Penelitian.....	33
3.4	Teknik Pengambilan Data	34
3.5	Perancangan Sistem.....	35
3.6	Analisis Kerja Sistem	44
BAB IV HASIL DAN PEMBAHASAN.....		49
4.1	Hasil Penelitian	49
4.2	Pembahasan.....	51
BAB V PENUTUP		53
5.1	Kesimpulan.....	53
5.2	Saran.....	53
DAFTAR PUSTAKA		54
LAMPIRAN.....		57

DAFTAR TABEL

Nomor	Halaman
Tabel 1 Contoh bentuk matriks $M \times N$ dari citra	8
Tabel 2 <i>State of the Art</i>	30
Tabel 3 Perhitungan dengan <i>confusion matrix</i>	50
Tabel 4 Performa sistem	50

DAFTAR GAMBAR

Nomor	Halaman
Gambar 1 Gambar logo mirip.....	6
Gambar 2 <i>Digital image</i>	9
Gambar 3 Sistem <i>computer vision</i> (Fisher, 2013)	10
Gambar 4 Contoh <i>flipping</i> (Pai, n.d.)	13
Gambar 5 Contoh <i>flipping</i> pada gedung	14
Gambar 6 Contoh <i>cropping</i>	15
Gambar 7 Contoh <i>rotation</i> (Pai, n.d.).....	16
Gambar 8 Contoh <i>color space transformations</i> dengan mengubah pencahayaan gambar (Mikołajczyk & Grochowski, 2018)	17
Gambar 9 Contoh <i>color space transformations</i> dengan menerapkan perubahan warna RGB (Wu et al., 2015).....	17
Gambar 10 Contoh <i>noise injection</i> (Pai, n.d.)	18
Gambar 11 Contoh <i>random erasing</i> (Zhong et al., 2020)	19
Gambar 12 Contoh <i>bounding boxes</i> 2D (Ambalina, 2019).....	21
Gambar 13 Contoh <i>bounding boxes</i> 3D (Ambalina, 2019).....	21
Gambar 14 Contoh <i>polygon</i> (Low, 2020).....	22
Gambar 15 Contoh <i>lines</i> (Low, 2020)	23
Gambar 16 Contoh <i>semantic segmentation</i> (Pokhrel, 2020).....	23
Gambar 17 Contoh <i>key-point</i> (Pokhrel, 2020).....	24
Gambar 18 Tampilan <i>Google Colaboratory</i>	27
Gambar 19 Kerangka pikir	31
Gambar 20 Tahapan penelitian.....	32
Gambar 21 Dokumen yang telah diunggah oleh dosen atau admin akademik....	34
Gambar 22 Ijazah dari website forlap.ristekdikti.go.id	35
Gambar 23 Ijazah untuk data latih dan data <i>testing</i>	35
Gambar 24 Desain alur sistem.....	36
Gambar 25 Augmentasi dengan rotasi.....	37
Gambar 26 <i>Flowchart</i> proses pelabelan gambar	38
Gambar 27 Pelabelan dengan menggunakan <code>labelImg.exe</code>	39
Gambar 28 File *.txt hasil pelabelan.....	39

Gambar 29 Perintah unduh data pelatihan.....	40
Gambar 30 Perintah unduh model YOLO	40
Gambar 31 Perintah pelatihan	40
Gambar 32 <i>Flowchart</i> proses <i>testing</i>	41
Gambar 33 Hasil deteksi.....	43
Gambar 34 <i>Confusion matrix</i>	44
Gambar 35 Contoh TP, FN, FP	46
Gambar 36 Metrik perhitungan nilai <i>Precision</i> , <i>Recall</i> , dan <i>F1-score</i>	47
Gambar 37 15 Perguruan tinggi untuk deteksi logo	49

DAFTAR LAMPIRAN

Nomor	Halaman
Lampiran 1 <i>Dataset</i>	58
Lampiran 2 Koleksi gambar ijazah dari website forlap.ristekdikti.go.id, dalam bentuk file *.jif dan *.pdf.....	63
Lampiran 3 <i>Source code</i>	65

BAB I

PENDAHULUAN

1.1 Latar Belakang

Logo pada saat ini yang menjadi bagian dari identitas yang melekat pada perusahaan. Identitas suatu perusahaan merupakan cerminan dari visi, misi dan ideal suatu perusahaan yang divisualisasikan ke dalam logo perusahaan (Martadi Martadi, 2002). Logo dapat berbentuk gambar (*logogram*) ataupun huruf (*logotype*) bahkan gambar dan huruf berbaaur menjadi satu.

Menurut Dendi (1999) dalam (Martadi Martadi, 2002), logo diartikan sebagai segala sesuatu yang berupa lambang, gambar, tulisan, angka, atau gabungan dari berbagai hal tersebut yang disandang oleh suatu produk, perusahaan, lembaga, organisasi, atau kegiatan, untuk mencirikan suatu eksistensinya agar dapat dibedakan dari produk atau merk lain.

Para ahli komunikasi mengungkapkan bahwa orang cenderung untuk mengingat gambar lebih kuat dari teks atau tulisan. Ketika orang melihat gambar yang menarik, mereka akan menyimpannya dalam ingatan lebih lama daripada tulisan.

Sebuah logo memiliki keunikan tersendiri dan memiliki representasi yang berbeda. Namun, ada beberapa logo yang sangat mirip dalam beberapa aspek, seperti bentuk atau pola yang sama, warna yang sama, jenis tulisan yang sama, dan lain-lain.

Logo biasanya digunakan dalam dokumen bisnis dan pemerintah sebagai pernyataan sumber dan kepemilikan dokumen. Salah satu dokumen yang memuat logo adalah ijazah. Ijazah merupakan dokumen yang dikeluarkan oleh suatu lembaga pendidikan sebagai tanda seseorang telah menyelesaikan jenjang pendidikan. Ijazah dapat digunakan untuk berbagai macam tujuan, seperti melamar pekerjaan, melanjutkan pendidikan, dan sebagainya.

Dalam lingkup pekerjaan, ijazah dibutuhkan sebagai berkas persyaratan untuk perubahan data pribadi seperti perubahan jenjang pendidikan, kenaikan pangkat, dan lain-lain. Yang kemudian berkas ijazah tersebut akan dicek dan didata

oleh pihak tertentu untuk diketahui asal lembaga pendidikan yang mengeluarkan ijazah tersebut.

Setelah observasi secara langsung terkait dengan pengecekan asal perguruan tinggi pada ijazah di Kopertis Wilayah IX, ijazah masih dicek secara satu per satu untuk mengetahui asal perguruan tinggi pada ijazah. Sehingga diperlukan sistem yang dapat mendeteksi asal perguruan tinggi pada ijazah dengan berdasarkan logo. Logo memungkinkan untuk menentukan sumber dokumen dengan cepat dan akurat (Le et al., 2014).

Seiring berkembangnya teknologi, begitu banyak pengembangan yang dilakukan untuk penelitian dalam menemukan dan mendeteksi tulisan, logo maupun objek tertentu dalam sebuah gambar dengan menggunakan *image processing*.

Pada penelitian sebelumnya oleh (Dixit & Shirdhonkar, 2016) membuat sistem untuk deteksi logo otomatis dan ekstraksi dari gambar dokumen menggunakan *Singular Value Decomposition (SVD)*. Penelitian selanjutnya oleh (Abdulmunim & Abass, 2017), membuat sistem untuk mendeteksi dan melakukan pencocokan logo pada dokumen berbahasa Arab dengan berbagai bentuk, ukuran, warna dan resolusi dengan menggunakan SURF (*Speeded Up Robust Features*). Kemudian, (Sharma et al., 2018) membuat sistem untuk mendeteksi tanda tangan dan logo pada sebuah dokumen menggunakan dua metode dari *Deep Convolutional Neural Networks (CNN)* yaitu *Faster R-CNN* dan *YOLOv2*.

Ada banyak tantangan dalam mendeteksi logo pada sebuah dokumen, seperti skala logo yang berbeda-beda, gambar dokumen yang buram (resolusi rendah), adanya perbedaan warna pada logo (dokumen scan asli dan fotokopi), rotasi (gambar dokumen miring), kemudian logo pada ijazah memiliki ukuran yang relatif kecil.

Untuk mendeteksi objek tertentu dalam sebuah gambar, sistem harus menemukan kecocokan titik fitur yang saling berhubungan antara titik fitur pada sebuah citra referensi dengan titik fitur pada citra target, dengan pengertian lain bagaimana menemukan keseragaman informasi terkait dengan ukuran, intensitas, rotasi, skala, *noise* dan *viewpoint* (Harahap, 2016).

Manusia sangat mudah untuk mengenali sebuah gambar karena otak manusia dengan mudah belajar dan melakukan pengenalan terhadap suatu objek dengan sangat akurat. Namun, mesin atau komputer belajar dan melakukan pengenalan terhadap objek dengan cara mengamati sebuah pola. Salah satu pendekatan yang dapat dilakukan adalah dengan *deep learning*. *Deep learning* termasuk dalam kategori *artificial intelligence* yang dapat berpikir layaknya manusia.

Penelitian ini menggunakan metode *You Only Look Once* versi 4 untuk mendeteksi logo pada dokumen berupa ijazah, dimana ijazah merupakan salah satu dokumen yang kompleks karena di dalamnya terdapat tanda tangan, logo, stempel, tulisan yang dicetak dengan ukuran dan jenis *font* yang berbeda (Shirdhonkar & Kokare, 2010). Dengan demikian penelitian ini mengambil judul “Deteksi Logo Dengan Menggunakan Metode *You Only Look Once* (YOLO)”.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang di atas, maka rumusan masalah dalam penelitian ini adalah:

1. Bagaimana mengimplementasikan metode YOLO untuk mendeteksi logo?
2. Berapa akurasi yang dihasilkan metode YOLO dalam mendeteksi logo dengan perhitungan *mean Average Precision* (mAP)?

1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah

1. Menerapkan metode YOLO untuk mendeteksi logo.
2. Untuk mengetahui berapa akurasi yang dihasilkan menggunakan metode YOLO dalam mendeteksi logo dengan perhitungan *mean Average Precision* (mAP).

1.4 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah diharapkan dapat memberikan pengetahuan tentang pengimplementasian metode YOLO untuk mendeteksi logo.

Dan dapat digunakan sebagai referensi untuk pengembangan sistem untuk mendeteksi objek yang berbeda.

1.5 Batasan Masalah

Batasan masalah penelitian adalah:

1. Data yang digunakan adalah gambar dokumen berupa ijazah perguruan tinggi dengan format *.JPG.
2. Metode yang digunakan adalah YOLO versi 4.

1.6 Sistematika Penulisan

Sistematika penulisan pada penelitian ini adalah:

BAB I PENDAHULUAN

Bab I berisi penjelasan tentang latar belakang yang mencakup apa peran dari sebuah logo, kendala dalam deteksi logo pada dokumen, penelitian yang terkait dengan deteksi logo. Terdapat rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah pada penelitian serta sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Bab II berisi penjelasan tentang teori-teori yang digunakan dalam penelitian dan kerangka pemikiran. Landasan teori merupakan suatu penjelasan tentang sumber acuan terbaru dari pustaka primer seperti buku, artikel, jurnal, prosiding dan tulisan asli lainnya untuk mengetahui perkembangan penelitian yang relevan dengan judul atau tema penelitian yang dilakukan dan juga sebagai arahan dalam memecahkan masalah yang diteliti. Dalam bab ini juga diuraikan tentang kerangka pikir tentang masalah, metode, pengukuran dan hasil dari penggunaan metode yang sesuai dengan objek dalam penelitian yang diusulkan.

BAB III METODOLOGI PENELITIAN

Bab III berisi penjelasan tentang tahapan penelitian, bagaimana pengembangan dan penerapan sistem yang diusulkan mengenai sistem untuk mendeteksi logo menggunakan metode YOLO, rancangan sistem dan uraian proses

validasi hasil dari metode yang diusulkan, serta analisa data pada hasil penerapan sistem yang diusulkan.

BAB IV HASIL DAN PEMBAHASAN

Pada bab IV ini menjelaskan tentang hasil dan pembahasan penelitian dari penelitian yang dilakukan. Hasil merupakan suatu penjelasan tentang data kuantitatif yang dikumpulkan dari lapangan sesuai dengan metodologi yang telah ditetapkan. Pembahasan merupakan suatu penjelasan tentang pengolahan data dan interprestasinya, baik dalam bentuk diskriptif ataupun penarikan inferensinya.

BAB V PENUTUP

Pada bab V ini berisi ringkasan temuan, rangkuman kesimpulan dan saran. Kesimpulan merupakan pernyataan secara general atau spesifik yang berisi hal-hal penting dan menjadi temuan penelitian yang bersumber pada hasil dan pembahasan. Saran merupakan pernyataan atau rekomendasi peneliti yang berisi hal-hal penting sebagaimana yang telah disampaikan.

BAB II

TINJAUAN PUSTAKA

2.1 Landasan Teori

2.1.1 Logo

Logo berasal dari bahasa Yunani yaitu *logos* yang berarti pikiran atau adalah penyajian atau tampilan nama, bentuk seragam, tulisan, atau ciri khas perusahaan secara visual (M . Linggar Anggoro, 2001).

Logo pada saat ini yang menjadi bagian dari identitas yang melekat pada perusahaan. Identitas suatu perusahaan merupakan cerminan dari visi, misi dan ideal suatu perusahaan yang divisualisasikan ke dalam logo perusahaan (Martadi Martadi, 2002). Logo dapat berbentuk gambar (*logogram*) ataupun huruf (*logotype*) bahkan gambar dan huruf berbaur menjadi satu.



Gambar 1 Gambar logo mirip

Menurut Dendi (1999) dalam (Martadi Martadi, 2002), logo diartikan sebagai segala sesuatu yang berupa lambang, gambar, tulisan, angka, atau gabungan dari berbagai hal tersebut yang disandang oleh suatu produk, perusahaan, lembaga, organisasi, atau kegiatan, untuk mencirikan suatu eksistensinya agar dapat dibedakan dari produk atau merk lain.

Sebuah logo memiliki keunikan tersendiri dan memiliki representasi yang berbeda. Namun, ada beberapa logo yang sangat mirip dalam beberapa aspek,

seperti bentuk atau pola yang sama, warna yang sama, jenis tulisan yang sama, dan lain-lain, seperti pada Gambar 1. Contohnya pada logo *Ford* dan *Carrier*, sama-sama memiliki bentuk yang sama yaitu *oval* (lonjong), kombinasi warna yang sama yaitu hitam dan putih, hal yang membedakan dari kedua logo tersebut adalah jenis *font*. Kemudian pada logo *Sun microsystems* dan *Columbia* sama-sama memiliki bentuk yang sama yaitu komponen yang terkandung di dalamnya membentuk sebuah logo menyerupai bentuk layang-layang, walaupun komponen yang terkandung di dalamnya itu berbeda. Kemudian pada logo gojek, *CircleCI*, memiliki bentuk yang hampir sama, yaitu berbentuk lingkaran tidak penuh dan memiliki komponen berbentuk bundar di dalamnya yang membedakan adalah ukuran dari komponen yang berada di dalam, posisi dari ujung logo yang berbeda, dan juga warna yang berbeda. Dan terakhir adalah logo *Gucci* dan *Channel*, memiliki komposisi warna yang sama yakni hitam, dan menyilangkan dua huruf yang sama yakni C untuk merk *Channel* dan G untuk merk *Gucci*.

Dibutuhkan dua proses dalam melakukan proses analisis logo pada sebuah gambar dokumen (Abdulmunim & Abass, 2017):

1. Deteksi logo : pada gambar dokumen akan diterapkan proses untuk mendeteksi logo.
2. Pengenalan logo : logo yang berhasil terdeteksi dari dokumen yang tersegmentasi, selanjutnya akan diklasifikasikan dan kemudian dimasukkan ke dalam kelasnya masing-masing.

2.1.2 Citra

Citra (*image*) adalah kombinasi antara titik, garis, bidang, dan warna untuk menciptakan suatu imitasi dari suatu objek-biasanya objek fisik atau manusia. Citra bisa berwujud gambar dua dimensi, seperti lukisan, foto, dan berwujud tiga dimensi, seperti patung.

Berdasarkan sinyal pembentuknya, citra dibedakan menjadi dua jenis yaitu citra analog dan citra digital. Citra analog merupakan citra yang terbentuk dari sinyal kontinyu. Nilai intensitas cahaya pada citra analog memiliki *range* antara 0 s.d ∞ . Alat akuisisi citra analog antara lain mata manusia dan kamera analog. Sedangkan citra digital merupakan citra yang terbentuk dari sinyal diskrit. Nilai

intensitas cahaya pada citra digital bergantung pada kedalaman *bit* yang menyusunnya.

Citra digital dapat dinyatakan dalam bentuk matriks $M \times N$, dimana:

M adalah jumlah baris; $0 \leq x \leq M-1$

N adalah jumlah kolom; $0 \leq x \leq N-1$

Contoh bentuk matriks dari citra dapat dilihat pada Tabel 1.

Tabel 1 Contoh bentuk matriks $M \times N$ dari citra

$f(x,y)$	$f(0,0)$	$f(0,1)$	$f(0,N-1)$
	$f(1,0)$	$f(1,1)$	$f(1,N-1)$

	$f(M-1,0)$	$f(M-1,1)$	$f(M-1, N-1)$

2.1.3 Pengolahan citra

Pengolahan citra adalah salah satu cabang dari ilmu informatika (komputer). Pengolahan citra berfokus pada usaha untuk melakukan transformasi suatu citra atau gambar menjadi citra lain dengan menggunakan teknik tertentu. Teknik-teknik pengolahan citra mentransformasikan citra menjadi citra lain. Jadi, masukannya adalah citra dan keluarannya juga citra, namun citra keluaran mempunyai kualitas lebih baik daripada citra masukan, dengan tujuan untuk menyempurnakan citra atau mendapatkan informasi yang berguna untuk diolah menjadi beberapa keputusan. Dalam operasi pemrosesan citra, operasi yang sering dilakukan dalam gambar *grayscale*. Gambar *grayscale* didapatkan dari pemrosesan gambar berwarna yang didekomposisi menjadi komponen merah (R), hijau (G), dan biru (B) yang diproses secara independen sebagai gambar *grayscale*. *Image processing* terbagi menjadi dalam 3 tingkatan (Vipin Tyagi, 2018).

1. *Low-level image processing*

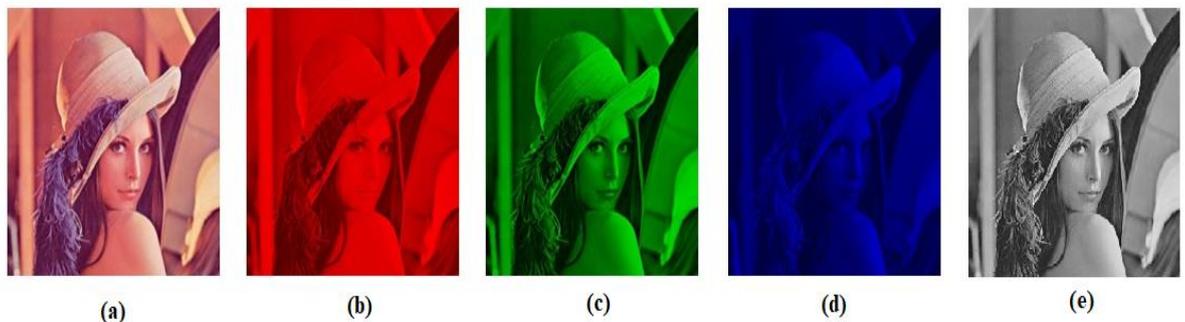
Merupakan operasi sederhana dalam pengolahan gambar dimana *input* dan *output* berupa gambar. Contoh: *contrast enhancement* dan *noise reduction*.

2. *Mid-level image processing*

Merupakan operasi pengolahan gambar yang melibatkan ekstraksi atribut dari gambar *input*. Contoh: *edges*, *contours* dan *regions*.

3. *High-level image processing*

Merupakan kategori yang melibatkan pemrosesan gambar kompleks yang terkait dengan analisis dan interpretasi konten dalam sebuah keadaan untuk pengambilan keputusan.

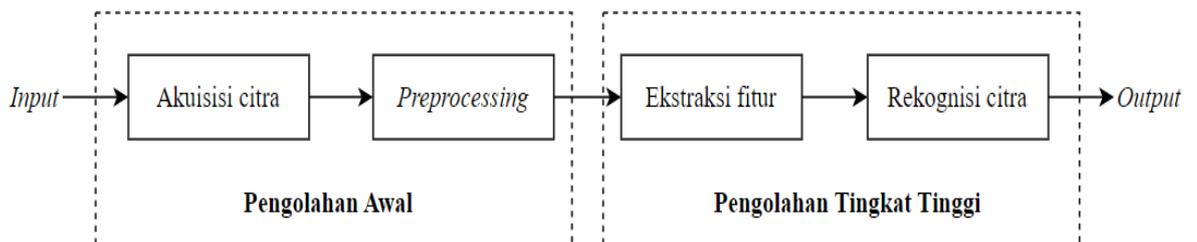


Gambar 2 *Digital image*; (a) Gambar berwarna, (b) Komponen merah dalam gambar berwarna, (c) Komponen hijau dalam gambar berwarna, (d) Komponen biru dalam gambar berwarna, (e) Gambar berwarna dikonversi dalam 8bit *grayscale*

Digital image merupakan fungsi dua dimensi $f(x,y)$ yang merupakan proyeksi dari bentuk tiga dimensi ke dalam bentuk dua dimensi dimana x dan y merupakan lokasi elemen gambar atau piksel yang berisikan nilai. Ketika nilai x , y dan intensitasnya berupa diskrit, maka gambar tersebut dapat dikategorikan sebagai *digital image*. Secara matematis, *digital image* adalah representasi matriks dari gambar dua dimensi menggunakan piksel. Setiap piksel diwakili oleh nilai numerik. Untuk gambar *grayscale*, hanya memiliki satu nilai dengan kisaran antara 0-255. Untuk gambar yang berwarna, memiliki tiga nilai yang mewakili *red* (R), *green* (G), dan *blue* (B) yang masing-masing memiliki kisaran nilai yang sama antara 0-255 pada Gambar 2. Jika suatu gambar hanya memiliki dua intensitas, gambar tersebut dikenal sebagai *binary image* (Vipin Tyagi, 2018).

2.1.4 *Computer vision*

Computer vision merupakan suatu bidang yang berhubungan dengan pencitraan otomatis dan pengolahan otomatis suatu citra berbasis komputer untuk mengestrak dan menginterpretasikan informasi. *Computer vision* bekerja dengan cara memproses data citra menggunakan kombinasi algoritma pengolahan citra dan kecerdasan buatan sehingga dapat menghasilkan informasi dan citra tersebut.



Gambar 3 Sistem *computer vision* (Fisher, 2013)

Computer vision dalam melaksanakan tugas-tugas membutuhkan proses yang kompleks. Biasanya, untuk setiap aplikasi yang diberikan, keseluruhan tugas tidak dapat dilaksanakan pada sebuah tahapan tunggal. *Computer vision* tertentu terdiri tahapan-tahapan seperti perolehan citra, *preprocessing*, pengekstraksian fitur, penyimpanan objek secara asosiatif, pengaksesan suatu basis pengetahuan, dan pengenalan (Fisher, 2013), seperti yang ditampilkan pada Gambar 3.

2.1.5 Pengenalan pola

Pengenalan pola (*pattern recognition*) merupakan bidang dalam pembelajaran mesin dan dapat diartikan sebagai “tindakan mengambil data mentah dan bertindak berdasarkan klasifikasi data”. Dengan demikian, pengenalan pola merupakan himpunan kaidah bagi pembelajaran diselia (*supervised learning*).

Ada beberapa definisi lain tentang pengenalan pola, di antaranya, penentuan suatu objek fisik atau kejadian ke dalam salah satu atau beberapa kategori, ilmu pengetahuan yang menitikberatkan pada deskripsi dan klasifikasi (pengenalan) dari suatu pengukuran.

2.1.6 Segmentasi citra

Segmentasi citra dianggap sebagai komponen utama pada setiap sistem *image processing*. Karena gambar terdiri dari beberapa *pixel*, maka dibagi menjadi beberapa segmen. Tujuan dari tahap ini untuk membuat analisis dan mengekstrak objek dari gambar digital lebih mudah, seperti mencari bentuk tertentu dan menemukan batas dari suatu objek.

2.1.7 Thresholding

Thresholding merupakan salah satu metode segmentasi citra yang memisahkan antara objek dengan *background* dalam suatu citra berdasarkan pada perbedaan tingkat kecerahannya atau gelap terangnya.

Pada operasi ini nilai piksel yang memenuhi syarat ambang batas dipetakan ke suatu nilai yang dikehendaki. Dalam hal ini syarat ambang batas dan nilai yang dikehendaki disesuaikan dengan kebutuhan. Operasi *thresholding* yang mempunyai ketentuan berikut: nilai intensitas *output* $f_o(x,y) = 0$, bila nilai intensitas *input* $f_i(x,y) \leq 0$, nilai $f_o(x,y) = T_1$ bila $T_1 < f_i(x,y) \leq T_2$, nilai $f_o(x,y) = T_2$ bila $T_2 < f_i(x,y) \leq T_3$, ..., nilai $f_o(x,y) = T_{n-1}$ bila $T_{n-1} < f_i(x,y) < T_n$, secara matematis dapat ditulis pada persamaan (1) berikut:

$$f_o(x,y) = \begin{cases} T_1, & f_i(x,y) \leq T_1 \\ T_2, & T_1 < f_i(x,y) \leq T_2 \\ T_3, & T_2 < f_i(x,y) \leq T_3 \\ \vdots & \vdots \\ T_n, & T_{n-1} < f_i(x,y) \leq T_n \end{cases} \quad (1)$$

Dalam hal ini $T_1, T_2, T_3, \dots, T_n$ adalah nilai-nilai yang dikehendaki, sedangkan $f_i(x,y) \leq T_1, T_1 < f_i(x,y) \leq T_2, T_2 < f_i(x,y) \leq T_3, \dots, T_{n-1} < f_i(x,y) \leq T_n$ adalah ambang batas yang disyaratkan. Nilai-nilai ambang $T_1, T_2, T_3, \dots, T_n$ dan syarat ambang batasnya ditentukan sesuai kebutuhan (T Sutojo, E Mulyanto, V Suhartono, OD Nurhayati, Wijanarto, 2009).

2.1.8 Augmentasi

Augmentasi adalah proses memodifikasi atau memanipulasi suatu gambar, sehingga gambar asli dalam bentuk standar akan diubah bentuk dan posisinya. Proses augmentasi adalah memodifikasi citra asli dari kumpulan data (*dataset*) yang sudah ada dengan cara membuat titik-titik data baru, misalnya dengan membalik, memutar citra asli, atau mengubah citra dengan berbagai pendekatan untuk membentuk citra baru (Rajcomar et al., 2020). Augmentasi data merupakan salah satu teknik yang dapat mengurangi *overfitting* dengan meningkatkan ukuran *dataset*. *Overfitting* mengacu pada fenomena ketika jaringan mempelajari fungsi dengan varian yang sangat tinggi seperti untuk memodelkan data pelatihan dengan sempurna (Shorten & Khoshgoftaar, 2019).

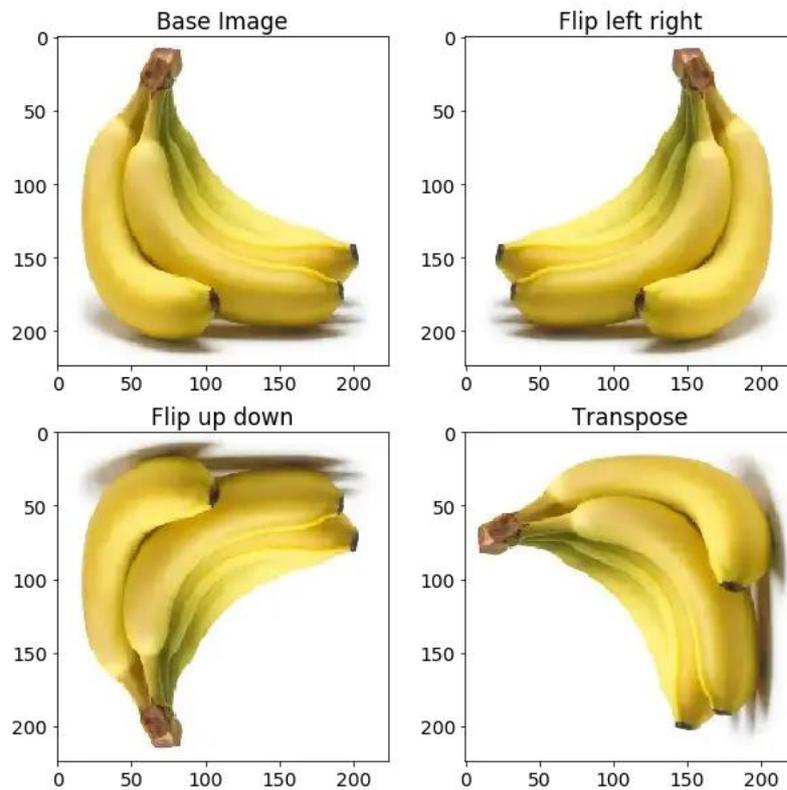
Augmentasi data bertujuan untuk memungkinkan mesin mempelajari dan mengenali gambar yang berbeda dan menggunakannya untuk mereproduksi data. Data biasanya diaugmentasi dengan melakukan transformasi pada data atau dengan kata lain membuat salinan dari sumber data tanpa mengubah label yang tertera pada setiap bagian dari data tersebut (Fadillah et al., 2021). Augmentasi data dapat meningkatkan kinerja model dan memperluas kumpulan *dataset* yang terbatas untuk memanfaatkan kemampuan *dataset* yang besar.

Beberapa teknik augmentasi data berdasarkan manipulasi gambar (Shorten & Khoshgoftaar, 2019), yaitu:

1. *Flipping*

Flipping adalah jenis augmentasi dengan untuk membalik posisi suatu citra secara vertikal atau horizontal, seperti yang ditunjukkan pada Gambar 4.

Dapat dilihat pada Gambar 4 yang menampilkan gambar pisang, gambar akan dibalik secara horizontal, dimana gambar pisang yang awalnya menghadap ke arah kanan, akan berubah menjadi pisang menghadap ke arah kiri. Selanjutnya gambar dibalik secara vertikal, dimana posisi tungkai pisang yang awalnya di atas, menjadi tungkai pisang berada di bagian bawah.



Gambar 4 Contoh *flipping* (Pai, n.d.)

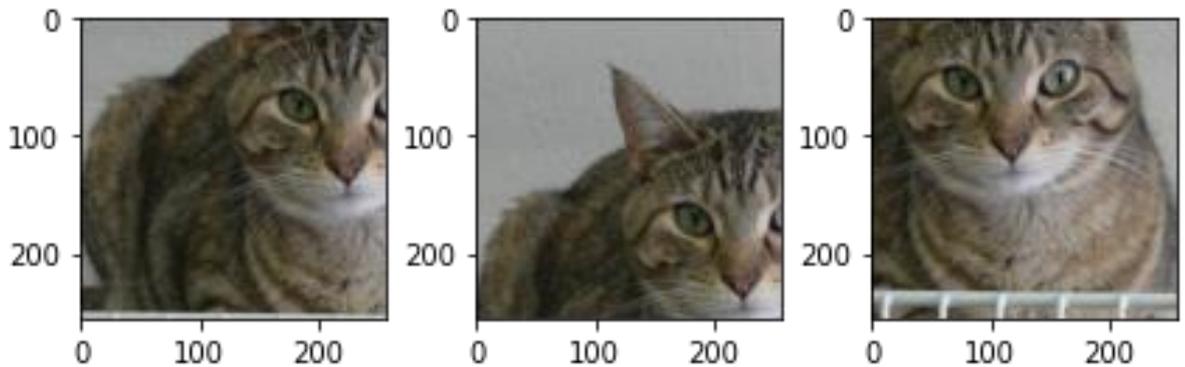
Namun, perlu diketahui bahwa tidak semua jenis citra/gambar boleh diterapkan teknik *flip*, dikarenakan dapat membuat citra menjadi tidak masuk akal. Contohnya pada Gambar 5, yaitu terdapat deretan dari beberapa gedung, diterapkan teknik *flipping* dengan membalik gambar gedung secara vertikal menghasilkan gambar dengan posisi atap gedung berada di bagian bawah, laut yang awalnya berada di bawah, menjadi di atas.



Gambar 5 Contoh *flipping* pada gedung

2. *Cropping*

Cropping (pemotongan) adalah salah satu jenis augmentasi dengan memotong bagian pada gambar, hasil dari *cropping* akan menghasilkan sebuah gambar yang tidak utuh, seperti pada Gambar 6.

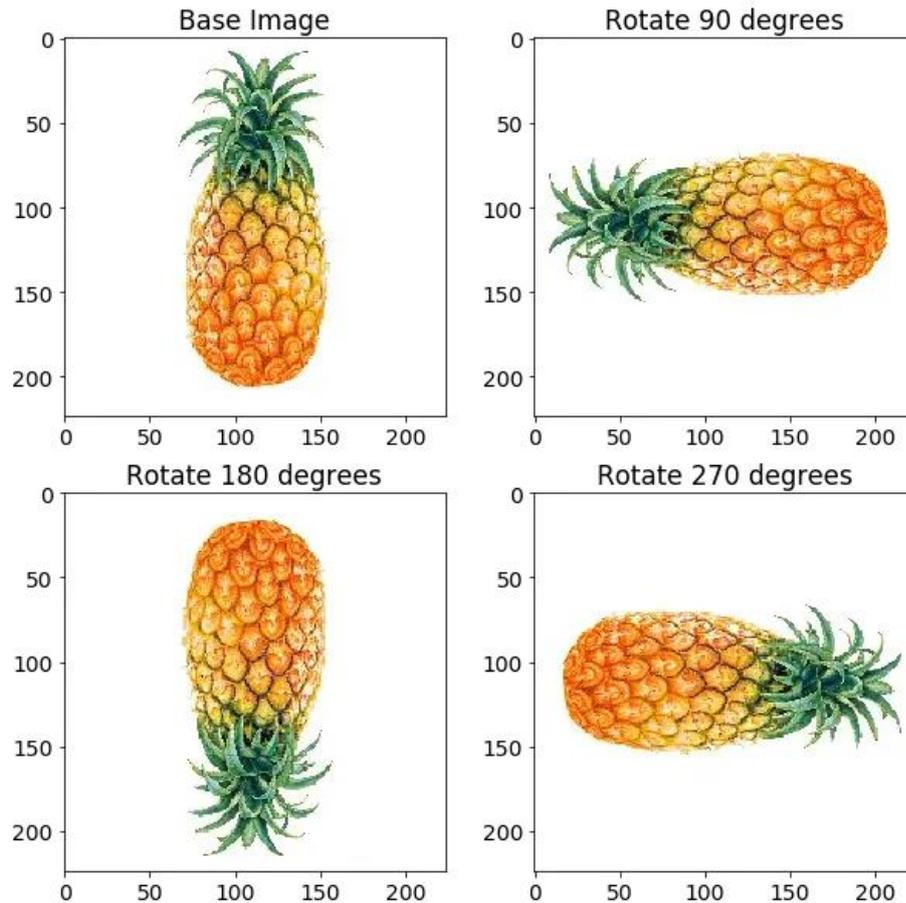


Gambar 6 Contoh *cropping*

Dapat dilihat gambar kucing dipotong dalam beberapa bagian, seperti *crop* $\frac{1}{2}$ gambar badan kucing, sehingga hanya terlihat bagian (badan, mata, hidung, mulut, leher, perut, paha, punggung) sebelah kiri. Kemudian *crop* $\frac{1}{2}$ gambar badan ditambah dengan *crop* bagian bawah gambar kucing, sehingga bagian bawah kucing tidak terlihat, dan hanya menyisakan tubuh bagian atas, dengan memperlihatkan bagian dari muka kucing (mata, hidung, telinga, mulut) serta punggung sebelah kiri. Kemudian *crop* yang hanya memperlihatkan badan kucing dari depan serta memperlihatkan seluruh bagian muka dari kucing baik kiri maupun kanan kecuali telinga.

3. *Rotation*

Rotation (rotasi) adalah jenis augmentasi yang dilakukan dengan memutar gambar ke kanan atau ke kiri pada sumbu antara 1° dan 359° , seperti pada Gambar 7. Dapat dilihat gambar nanas yang diubah rotasinya ke 90 derajat, menghasilkan gambar nanas yang berbeda dari gambar semula yaitu bagian kepala nanas berpindah ke kiri, kemudian mengubah rotasi 180 derajat, menghasilkan posisi kepala nanas berada di bawah. Selanjutnya mengubah gambar dengan rotasi 270 derajat, menghasilkan citra baru yaitu bagian kepala nanas bergeser ke sebelah kanan.

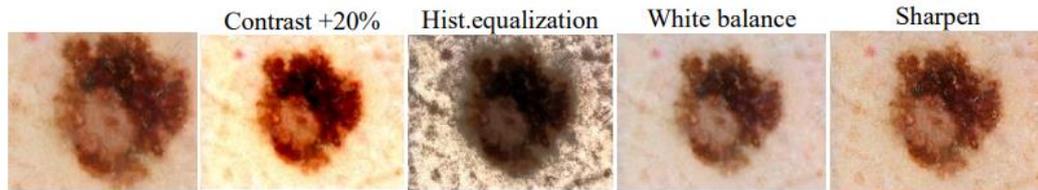


Gambar 7 Contoh *rotation* (Pai, n.d.)

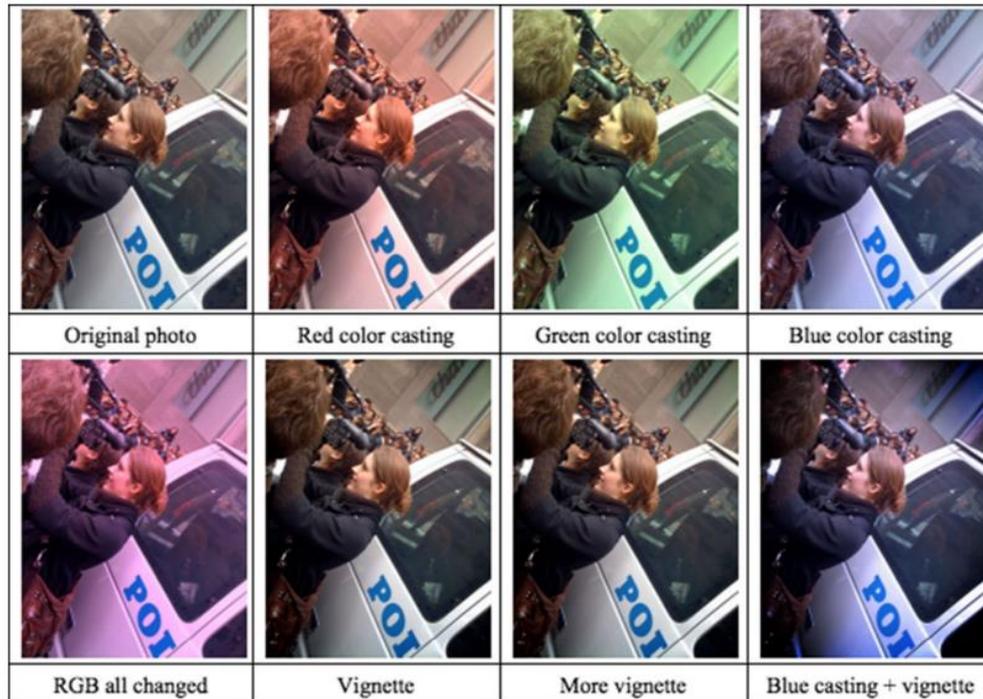
4. *Color space transformations*

Merupakan teknik augmentasi dengan cara mengubah warna dari citra, seperti pada Gambar 8 dan Gambar 9.

Pada Gambar 8 memperlihatkan citra luka pada kulit. Pada citra ini diterapkan peningkatan kontras atau kecerahan sebanyak +20% menghasilkan gambar citra yang lebih cerah dari citra asli. Kemudian diterapkan *histogram equalization* yang berfungsi untuk menyamakan warna kulit dengan warna luka pada kulit tersebut. Selanjutnya diterapkan *white balance* pada citra yang berfungsi untuk menyeimbangkan warna antar kulit dan luka dengan kalibrasi warna putih. Selanjutnya adalah mengatur ketajaman pada citra, sehingga gambar akan menjadi tampak lebih tajam atau jelas.



Gambar 8 Contoh *color space transformations* dengan mengubah pencahayaan gambar (Mikołajczyk & Grochowski, 2018)



Gambar 9 Contoh *color space transformations* dengan menerapkan perubahan warna RGB (Wu et al., 2015)

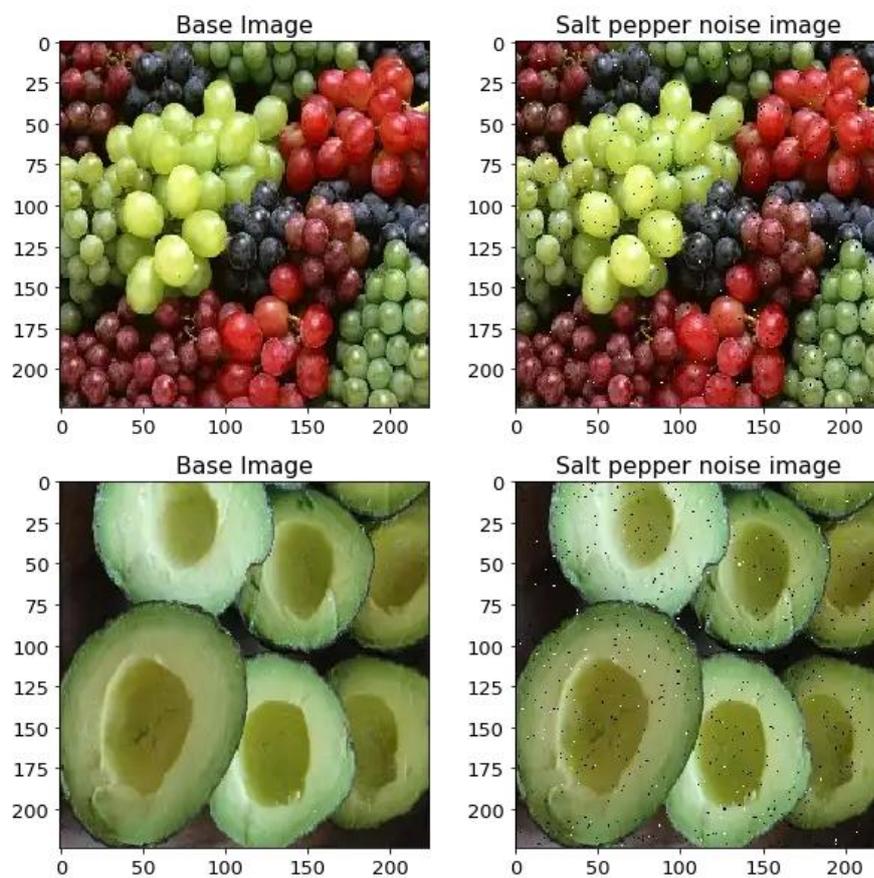
Pada Gambar 9, dapat dilihat gambar bagian dari mobil, wanita, dan kumpulan orang. Diberikan warna merah yang menghasilkan gambar cenderung berwarna merah, warna hijau akan membuat gambar cenderung berwarna hijau, warna biru akan membuat gambar cenderung berwarna biru. Kemudian menggabungkan ketiga warna (merah, hijau, biru). Selanjutnya memberikan efek *vignette*, yang berfungsi untuk membuat pinggiran gambar menjadi gelap dibandingkan dengan pusat gambar. Kemudian mengkombinasikan pemberian warnaa biru dan efek *vignette* pada gambar.

Adapun kelemahan dari *color space transformations* adalah dapat meningkatkan ukuran memori dari setiap data gambar, membutuhkan

waktu pelatihan yang lebih lama. Transformasi warna juga dapat membuang informasi warna yang penting pada citra asli. Kemudian peningkatan kecerahan dapat membuat citra asli menjadi tidak terlihat atau dikenali.

5. *Noise injection*

Merupakan teknik agumentasi dengan menambahkan *noise* pada gambar, seperti pada Gambar 10.

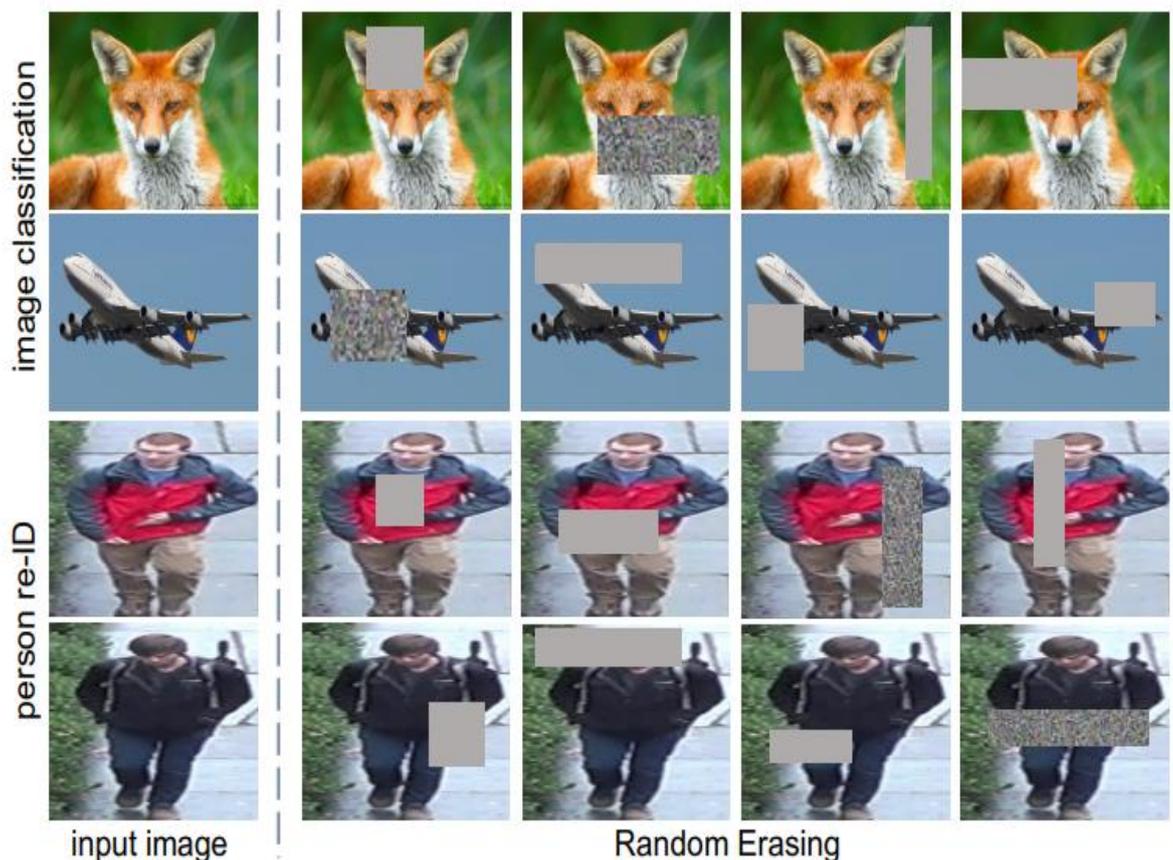


Gambar 10 Contoh *noise injection* (Pai, n.d.)

Pada Gambar 10 ditampilkan dua citra yaitu gambar beberapa jenis anggur dan gambar alpukat, pada kedua gambar tersebut diberikan *noise* berupa titik-titik berwarna hitam dan putih pada keseluruhan citra.

6. *Random erasing*

Random erasing (penghapusan acak) merupakan teknik augmentasi yang dikembangkan oleh (Zhong et al., 2020). Penghapusan acak secara acak memilih bidang persegi panjang dalam gambar dan menghapus pikselnya dengan nilai acak. Teknik ini secara khusus dirancang untuk mengatasi tantangan pengenalan gambar akibat oklusi. Oklusi mengacu pada ketika beberapa bagian objek tidak jelas. Penghapusan acak ini akan memaksa model mempelajari lebih banyak fitur deskriptif tentang suatu gambar, mencegahnya beralih ke fitur visual tertentu dalam gambar, seperti pada Gambar 11.



Gambar 11 Contoh *random erasing* (Zhong et al., 2020)

Pada Gambar 11 menampilkan empat jenis gambar yaitu rubah, pesawat, lelaki dengan jaket kombinasi warna merah dan abu-abu, dan lelaki yang memakai jaket berwarna hitam. Penghapusan dilakukan secara acak pada keempat gambar tersebut. Contohnya pada citra rubah,

menghapus sebagian dari bagian kepala dan telinga, menghapus bagian mulut dan batang hidung, menghapus ujung telinga bagian kanan, dan juga menghapus sebagian dari kening, telinga, dan mata sebelah kiri. Kemudian pada citra pesawat, dilakukan penghapusan pada bagian bagian tengah badan pesawat, moncong pesawat, sayap kiri dan juga sayap kanan pesawat. Selanjutnya pada citra laki-laki dengan jaket kombinasi warna merah dan abu-abu dilakukan penghapusan pada bagian dada, perut, bahu hingga betis kanan, kemudian ujung kepala hingga paha sebelah kiri. Dan pada citra laki-laki yang menggunakan jaket berwarna hitam dilakukan penghapusan pada bagian dada sampai paha, kepala, lutut, bagian perut, ujung tangan dan paha.

Cropping dan *flipping* merupakan dua transformasi yang paling sering digunakan dalam proses augmentasi data (Krizhevsky et al., 2017).

2.1.9 Anotasi gambar

Anotasi gambar adalah teknik yang digunakan untuk membuat data pelatihan untuk visi komputer. Gambar beranotasi dibutuhkan untuk melatih algoritma *machine learning* yang akan digunakan untuk belajar mengenali objek yang diajarkan. Anotasi dalam *machine learning* pada dasarnya adalah proses pelabelan data dalam berbagai media, seperti gambar, teks, atau video (Low, 2020).

Terdapat beberapa macam anotasi gambar yang biasa digunakan untuk visi komputer, seperti *bounding boxes*, *polygon*, *lines*, *semantic segmentation*, *key-point*.

1. *Bounding boxes*

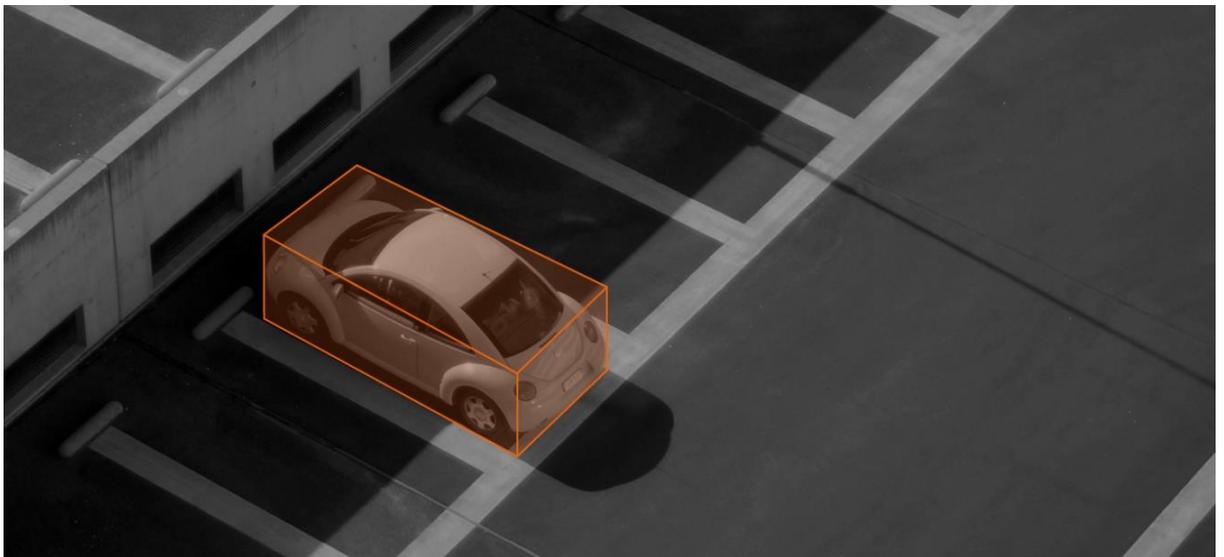
Bentuk anotasi ini mengharuskan pemberi label menggambar kotak sedekat mungkin dengan tepi objek utama di dalam gambar, seperti pada Gambar 12 dan Gambar 13.

Pada Gambar 12 dapat dilihat pemberian kotak pembatas berbentuk 2 dimensi pada setiap mobil dengan memperhatikan batas tepi dari mobil, sehingga objek tersebut dapat dikenali sebagai mobil.

Sama dengan gambar 12, pada Gambar 13 diberikan kotak pembatas berbentuk 3 dimensi, sehingga terlihat gambar mobil diletakkan di dalam sebuah *box*, dengan mengambil keseluruhan dari bagian mobil.



Gambar 12 Contoh *bounding boxes* 2D (Ambalina, 2019)



Gambar 13 Contoh *bounding boxes* 3D (Ambalina, 2019)

2. Polygon

Biasanya digunakan untuk anotasi yang lebih tepat untuk item yang bentuknya tidak beraturan.



Gambar 14 Contoh *polygon* (Low, 2020)

Di Gambar 14, berjejer tiga jenis tanaman yang berbeda. Pemberian label diberikan dengan cara mengikuti bentuk dari daun setiap tanaman.

3. *Lines*

Digunakan untuk menggambar batas di suatu wilayah gambar, biasanya digunakan ketika bagian yang perlu digambarkan terlalu kecil atau tipis dan tidak dapat dicapai dengan kotak pembatas. Contoh dari penggunaan *lines* dapat dilihat pada Gambar 15 yang memperlihatkan deretan mobil dengan berbagai jenis seperti mobil sedan, truk, dan sebagainya. Gambar diambil dari dekat menghasilkan gambar mobil terlihat besar dan gambar mobil di belakang terlihat kecil, sehingga sulit untuk memberikan kotak pembatas pada objek tersebut. Sehingga jenis pelabelan yang digunakan pelabelan bentuk *lines* (garis). *Line* hanya perlu diletakkan pada deretan mobil dari depan ke belakang (ujung jalan).



Gambar 15 Contoh *lines* (Low, 2020)

4. *Semantic segmentation*

Untuk memisahkan gambar menjadi beberapa bagian dan mengklasifikasikan setiap piksel di setiap segmen ke label kelas yang sesuai dari apa yang diwakilinya, seperti pada Gambar 16.



Gambar 16 Contoh *semantic segmentation* (Pohrel, 2020)

Di Gambar 16, pada dilihat pemberian label diberikan dengan cara memisahkan setiap bagian pada gambar dengan memberikan jenis warna yang berbeda. Misalnya, gambar mobil dipisahkan menjadi bagian berwarna biru, kemudian jalan dipisahkan menjadi bagian berwarna merah jambu, rambu lalu lintas dipisahkan menjadi bagian berwarna kuning.

5. *Key-point*



Gambar 17 Contoh *key-point* (Pokhrel, 2020)

Key-point digunakan untuk mendeteksi objek kecil dan variasi bentuk dengan membuat titik di seluruh gambar, seperti pada Gambar 17, diperlihatkan sekumpulan orang yang sedang duduk. Pada citra setiap orang tersebut ada diberikan titik-titik pada pada bagian tubuh. Seperti titik pada bagian kedua mata, kedua bahu, kedua siku, kedua pergelangan tangan, kedua ujung jari, pinggang. Semua titik-titik tersebut akan saling berhubungan dan akan membentuk sebuah kumpulan titik berbentuk “rangka” dari manusia.

2.1.10 YOLO

You Only Look Once (YOLO) adalah salah satu metode dalam *deep learning* deteksi objek tercepat saat ini. Jika biasanya untuk melakukan proses klasifikasi objek, algoritma yang dibuat adalah dengan menggunakan algoritma klasifikasi seperti *VGGNet* atau *Inception* yang dijalankan pada setiap bagian kecil dari citra menggunakan *sliding window*. Proses ini akan berhasil mendeteksi objek akan tetapi sistem akan sangat lambat, karena algoritma klasifikasi tersebut akan dijalankan berulang-ulang kali.

YOLO pertama kali diciptakan oleh Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi. YOLO membagi gambar atau video yang diinput menjadi $S \times S$ *grid*. Jika titik tengah koordinat pada GT (*Ground Truth*) suatu objek jatuh ke dalam *grid*, maka *grid* tersebut bertanggung jawab untuk mendeteksi suatu objek.

YOLO melakukan pendekatan yang sangat berbeda dengan algoritma lain. YOLO akan melakukan *scanning* pada citra hanya sekali dengan cara yang efisien. YOLO menggunakan *convolutional neural networks* untuk mendeteksi objek. Meskipun bukan algoritma yang terakurat, akan tetapi metode ini merupakan pilihan yang baik untuk melakukan deteksi objek secara *real-time*, tanpa harus kehilangan banyak akurasi.

Jika dibandingkan dengan algoritma *recognition* (pengenalan), metode pendeteksi objek ini tidak hanya memprediksi jenis kelas objek akan tetapi juga mendeteksi lokasi dari objek tersebut. Jadi, metode ini tidak hanya mengklasifikasi sebuah *input* citra ke dalam suatu kategori, akan tetapi juga bisa mendeteksi berbagai objek yang berada di dalam citra. Metode ini mengaplikasikan *single neural network* ke sebuah citra secara keseluruhan. Artinya, *network* tersebut membagi citra ke berbagai *regions* dan memprediksi *bounding boxes* dan probabilitas tiap *region*. Setiap *bounding box* yang ada dihitung berdasarkan probabilitas prediksinya.

YOLO versi 4 adalah metode deteksi objek generasi ke empat dari YOLO yang dikembangkan oleh Alexey Bochkovskiy, Chien-Yao Wang, dan Hong-Yuan Mark Liao dan dirilis pada 24 April 2020.

Pada versi ini, YOLO mengalami improvisasi dengan berbagai penyesuaian dan memiliki kemampuan untuk mendeteksi objek dengan cepat dan akurat dibandingkan dengan versi sebelumnya.

YOLO versi 4 berjalan dua kali lebih cepat dibandingkan *EfficientDet* dengan performa yang sama dan meningkatkan AP dan FPS sebesar 10% dan 12% dibandingkan YOLOv3 (Chethan Kumar et al., 2020). YOLO versi 4 berhasil mendeteksi beberapa kelas objek dalam sebuah video dengan mendeteksi setiap objek di setiap *frame* demi *frame* hingga seluruh *frame* video tertutup, kelas tersebut dilambangkan dalam *bounding box*.

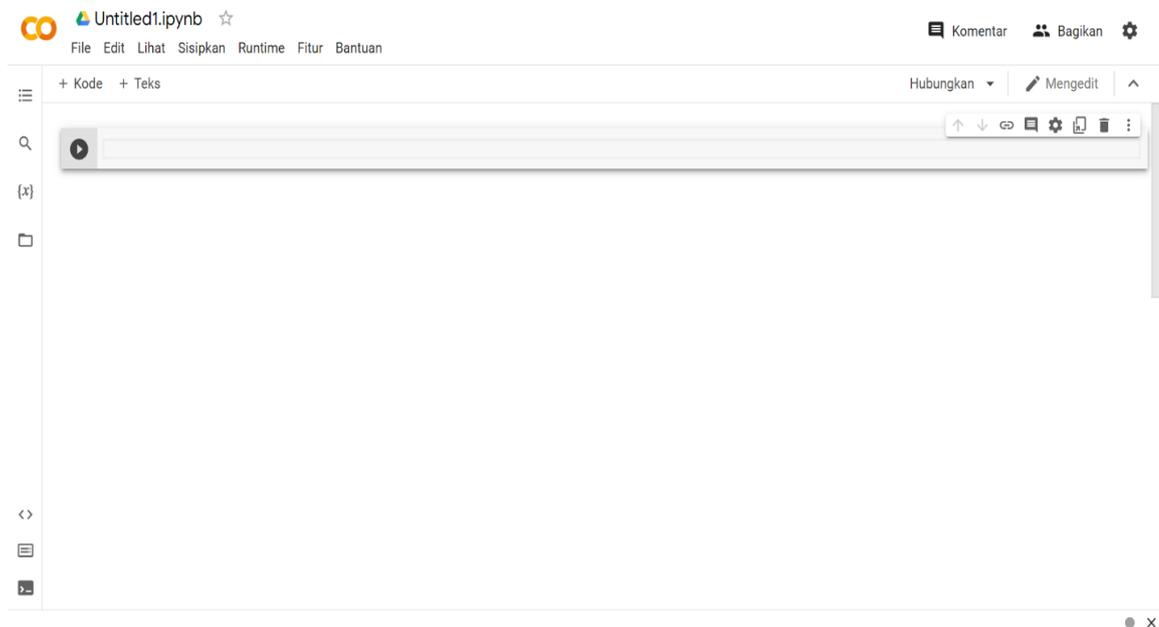
2.1.11 Google colabatory

Google Colab adalah salah satu produk Google berbasis *cloud* yang bisa digunakan secara gratis. Google Colab dibuat khusus untuk membantu para peneliti dalam mengolah data, khususnya di bidang *machine learning*. *Tools* Colab ini menyediakan layanan GPU gratis kepada pengguna sebagai *backend* komputasi. Dengan Google Colab, pengguna dapat mengakses sumber daya yang *powerful* melalui *browser* secara bebas selama 12 jam untuk mengeksekusi kode dan membangun aplikasi menggunakan berbagai *library* yang tersedia seperti *Numpy*, *OpenCV*, *Keras*, *TensorFlow*, dan lain-lain.

Google Colab adalah *coding environment* bahasa pemrograman *Python* dengan format “*notebook*” memiliki tampilan yang mirip dengan *Jupyter Notebook* (Gambar 18) dan tidak memerlukan pengaturan terlebih dahulu sebelum digunakan dan berjalan sepenuhnya pada *cloud* dengan memanfaatkan media penyimpanan Google Drive.

Selain itu, terdapat beberapa perintah yang bisa digunakan di Google Colab seperti perintah *git*, *pip* dan *sed*. *Git* merupakan fasilitas *version control system* yang banyak digunakan oleh programmer untuk berbagi paket program dan membangun aplikasi secara bersama-sama dengan menggunakan perintah dasar seperti *git init* (membuat *repository* atau tempat penyimpanan paket program), *git add* (menambah file baru pada *repository*), *git clone* (menyalin *repository*), dan lain-lain.

Pada Gambar 18, memperlihatkan tampilan dari Google Colab yang terdiri dari beberapa menu *File*, *Edit*, *Lihat*, *Sisipkan*, *Runtime*, *Fitur* dan *Bantuan*.



Gambar 18 Tampilan *Google Colaboratory*

Pada menu *File*, pengguna dapat menyimpan data, mengupload data baik dari komputer maupun dari penyimpanan berbasis *cloud*, mengunduh kode yang telah dibuat, dan tugas lainnya. Kemudian di menu *Edit*, pengguna dapat menghapus, mengubah *cell*, dan melakukan pengaturan pada *notebook*. Di menu *Lihat*, pengguna dapat mengatur konten apa saja yang ingin ditampilkan seperti *Table of Content*, *Notebook Info*, dan lain-lain. Kemudian di menu *Runtime*, pengguna dapat melakukan *run* pada kode yang telah dibuat, ataupun memberhentikan proses *run* yang telah dijalankan. Pada menu *Fitur*, pengguna dapat mengatur *shortcut* yang dapat digunakan saat menggunakan *Jupyter notebook* di *Google Colab*. Sedangkan di menu *Bantuan*, pengguna dapat melaporkan jika terjadi *bug* saat menggunakan fitur di *Google Colab* maupun mencari pertanyaan yang umum ditanyakan oleh pengguna lain terkait dengan *Google Colab*.

2.2 Penelitian Terkait

Beberapa penelitian terkait pendeteksian logo diantaranya sebagai berikut:

1. (Dixit & Shirdhonkar, 2016), membuat sistem untuk deteksi logo otomatis dan ekstraksi dari gambar dokumen menggunakan *Singular Value Decomposition* (SVD). Untuk menguji metode yang digunakan *dataset* diambil dari Tobacco-800. Kemudian dari hasil pengujian pada berbagai kategori logo mencapai tingkat deteksi logo rata-rata 89,52%.
2. (Abdulmunim & Abass, 2017), membuat sistem untuk mendeteksi dan melakukan pencocokan logo pada dokumen berbahasa Arab dengan berbagai bentuk, ukuran, warna dan resolusi dengan menggunakan SURF (*Speeded Up Robust Features*). Digunakan dua jenis fitur dalam melakukan pencocokan logo, yaitu fitur berbasis wilayah yang mewakili fitur global dari objek, dimana fitur ini akan membedakan jenis logo berdasarkan bentuk eksteriornya sehingga dapat mengurangi ruang pencarian *dataset*. Kemudian fitur kedua adalah deteksi fitur lokal dengan menggunakan SURF. Hasil penelitian menunjukkan performa sistem dalam pencocokan logo adalah *recall* 97%, presisi 99%, *F-score* 98%.
3. (Karaoglu et al., 2017), membuat sistem deteksi teks untuk klasifikasi *fine-grained object* pada *dataset* yang berasal dari ICDAR03, ICDAR13, Con-Text and Flickr-logo dengan cara menggabungkan isyarat tekstual dan visual untuk membedakan objek dengan menggunakan mesin OCR i.e ABBYY. Isyarat tekstual dan visual yang diekstraksi digabungkan untuk klasifikasi *fine-grained* dan pengambilan logo. Pada penelitian ini, metode yang diusulkan pada gambar yang diuji adalah dengan mendeteksi latar belakang untuk menyimpulkan lokasi teks, gambar yang diuji. Wilayah yang tersisa setelah menghilangkan latar belakang dianggap sebagai wilayah dari teks.
4. (Sharma et al., 2018), membuat sistem untuk mendeteksi tanda tangan dan logo pada sebuah dokumen menggunakan dua metode dari *Deep Convolutional Neural Networks* (CNN) yaitu *Faster R-CNN* dan YOLOv2. *Dataset* yang digunakan untuk sistem ini diambil dari Tobacco-800 sebanyak 1290 dokumen. Hasil penelitian menunjukkan metode *Deep Convolutional Neural Networks* (CNN) yang diusulkan mencapai 89,6% mAP. *Deep CNN* memiliki potensi dalam mendeteksi gambar dokumen secara *real time*, dan memiliki kinerja yang baik dalam mendeteksi tanda tangan maupun logo pada dokumen yang memiliki *noise*.

5. (Sarwo et al., 2019), membuat sistem untuk mendeteksi logo dan nama merek pada perusahaan dengan mengusulkan model varian ResNet baru yaitu menggunakan ResNet-50. Hasil empiris menunjukkan Model 2 mencapai 0,408 mAP, akurasi rata-rata terbaik dengan waktu pelatihan 1,41 jam. Sebaliknya, dengan menggunakan model Resnet50 mencapai akurasi rata-rata nilai *mean Average Precision* adalah 0,556 dengan waktu pelatihan 1,91 jam. Pengujian deteksi Model 2 yang diusulkan adalah 23,47 fps, sedangkan pengujian deteksi model Resnet50 adalah 29,33 Fps.

2.3 State of the Art

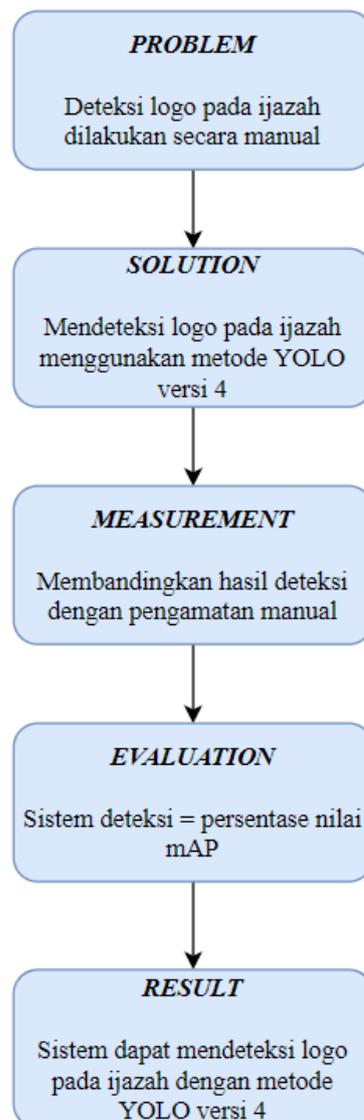
Penelitian sebelumnya oleh Dixit & Shirdhonkar menggunakan *Singular Value Decomposition* (SVD) untuk deteksi logo otomatis dan ekstraksi dari gambar dokumen. Matheel E. Abdulmunim & Haithem K. Abass menggunakan SURF (*Speeded Up Robust Features*) untuk mendeteksi dan melakukan pencocokan logo pada dokumen berbahasa Arab. Karaoglu et al mendeteksi teks untuk klasifikasi *fine-grained object* dengan menggunakan mesin OCR (*Optical Character Recognition*) dengan cara menggabungkan isyarat tekstual dan visual untuk membedakan objek. Nabin Sharma et al menggunakan *Deep Convolutional Neural Networks* (CNN) untuk mendeteksi tangan tangan dan logo pada gambar dokumen. Sarwo et al menggunakan ResNet-50 untuk mendeteksi logo dan nama merek pada perusahaan.

Tabel 2 *State of the Art*

No	Judul	Penerbit, Penulis & Tahun	Metode	Hasil
1.	<i>Automatic Logo Detection and Extraction using Singular Value Decomposition</i> (10.1109/ICCSP.2016.7754252)	<i>International Conference on Communication and Signal Processing</i> , Umesh D. Dixit, et al, 2016	Menggunakan <i>Singular Value Decomposition</i> (SVD)	Hasil pengujian pada berbagai kategori logo mencapai tingkat deteksi logo rata-rata 89,52%
2.	<i>Logo Matching in Arabic Documents Using Region Based Features and SURF Descriptor</i> (10.1109/NTICT.2017.7976128)	<i>2017 Annual Conference on New Trends in Information & Communications Technology Applications</i> (NTICT), Matheel E. Abdulmunim & Haithem K. Abass, 2017	Menggunakan SURF (<i>Speeded Up Robust Features</i>)	Hasil penelitian menunjukkan performa sistem dalam pencocokan logo adalah <i>recall</i> 97%, presisi 99%, <i>F-score</i> 98%.
3.	<i>Con-Text: Text Detection for Fine-grained Object Classification</i> (10.1109/TIP.2017.2707805)	<i>IEEE Transactions on Image Processing</i> , Sezer Karaoglu, et al, 2017	Menggunakan mesin OCR i.e ABBYY	Berdasarkan penelitian terkait, hasil untuk klasifikasi <i>fine-grained</i> meningkat dari 39,0% menjadi 70,7%, sedangkan pada pengambilan logo meningkat dari 54,8% menjadi 57,4%
4.	<i>Signature and Logo Detection using Deep CNN for Document Image Retrieval</i> (10.1109/ICFHR-2018.2018.00079)	<i>2018 16th International Conference on Frontiers in Handwriting Recognition</i> (ICFHR), Nabin Sharma, et al, 2018	Menggunakan <i>Faster R-CNN</i> dan YOLOv2	Hasil penelitian menunjukkan metode <i>Deep Convolutional Neural Networks</i> (CNN) yang diusulkan mencapai 89,6% mAP.
5.	<i>Logo detection and brand recognition with one-stage logo detection framework and simplified resnet50 backbone</i> (10.1109/AIT49014.2019.9144794)	<i>2019 International Congress on Applied Information Technology</i> (AIT), Sarwo et al, 2019	Menggunakan ResNet-50	Hasil menunjukkan akurasi rata-rata nilai <i>mean Average Precision</i> adalah 0,556 dengan waktu pelatihan 1,91 jam

2.4 Kerangka Pikir

Kerangka pikir akan ditampilkan pada Gambar 19, yang menjelaskan alur penelitian yang dilakukan. Pada tahap pertama menjelaskan permasalahan yang ada yaitu deteksi logo pada dokumen ijazah masih dilakukan secara manual. Solusi yang ditawarkan adalah sebuah sistem yang dapat mendeteksi logo pada ijazah secara otomatis dengan menggunakan metode YOLO versi 4, dimana sistem dapat mendeteksi asal perguruan tinggi berdasarkan logo yang dideteksi pada ijazah. Pengujian pada sistem dilakukan dengan membandingkan hasil deteksi dengan pengamatan manual. Adapun evaluasi yang dilakukan dengan menghitung *mean Average Precision* (mAP).



Gambar 19 Kerangka pikir