

SKRIPSI

**SISTEM DETEKSI JARAK RELATIF PADA AUTONOMOUS
SHIP DENGAN STEREOSCOPIIC MEASUREMENT**

Disusun dan diajukan oleh:

**DANDY GARDA DIRGANTARA
D121 18 1506**



**PROGRAM STUDI SARJANA TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS HASANUDDIN
GOWA
2023**

LEMBAR PENGESAHAN SKRIPSI

SISTEM DETEKSI JARAK RELATIF PADA AUTONOMOUS SHIP DENGAN STEREOSCOPIC MEASUREMENT

Disusun dan diajukan oleh


DANDY GARDA DIRGANTARA
D121181506

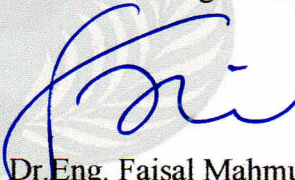
Telah dipertahankan di hadapan Panitia Ujian yang dibentuk dalam rangka Penyelesaian Studi Program Sarjana Program Studi Teknik Informatika Fakultas Teknik Universitas Hasanuddin Pada tanggal 08 Februari 2023 dan dinyatakan telah memenuhi syarat kelulusan

Menyetujui,


Pembimbing Utama,

Pembimbing Pendamping,


Prof. Dr. Indrabayu, ST., MT.,
M.Bus.Sys., IPM., ASEAN Eng.
NIP. 19750716 200212 1 004


Dr. Eng. Faisal Mahmuddin, ST.,
M. Inf. Tech., M. Eng.
NIP. 19810211 200501 1 003

Ketua Program Studi,


Prof. Dr. Indrabayu, ST., MT., M. Bus. Sys., IPM., ASEAN Eng.
NIP. 19750716 200212 1 004



PERNYATAAN KEASLIAN

Yang bertanda tangan dibawah ini ;

Nama : Dandy Garda Dirgantara
NIM : D121181506
Program Studi : Teknik Informatika
Jenjang : S1

Menyatakan dengan ini bahwa karya tulisan saya berjudul

Sistem Deteksi Jarak Relatif Pada Autonomous Ship Dengan Stereoscopic Measurement

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, 16 Februari 2023

Yang Menyatakan



Dandy Garda Dirgantara

ABSTRAK

DANDY GARDA DIRGANTARA. *Sistem Deteksi Jarak Relatif Pada Autonomous Ship Dengan Stereoscopic Measurement* (dibimbing oleh Prof. Dr. Indrabayu, ST., MT., M.Bus.Sys., IPM., ASEAN Eng. dan Dr.Eng. Faisal Mahmuddin, ST., M.Inf.Tech., M.Eng)

Teknologi Autonomous Ship atau kapal tanpa awak merupakan teknologi yang dapat mengoperasikan kapal tanpa memerlukan input dari manusia, dengan bekerja menggunakan kecerdasan buatan (Artificial Intelligence) untuk melihat keadaan sekitar kapal. Dalam pengembangan Autonomous Ship, rintangan yang harus dihadapi adalah seberapa dekat kapal dengan objek depannya, untuk itu diperlukan sistem yang dapat mendeteksi jarak relatif secara real-time. Dalam penelitian ini, akan menggunakan metode Computer Vision yang dapat memberikan informasi terkait objek sekitar seperti perahu nelayan (fishing boat) dan digabungkan dalam perhitungan Stereoscopic Measurement dengan kamera stereo yaitu ZED Stereo Camera, sehingga jarak antara perahu dapat diketahui secara langsung. Dan dalam penelitian ini dilakukan pendeteksian jarak antara perahu dari 7 hingga 20 meter. Hasil dari penelitian menunjukkan bahwa penggunaan metode YOLOv5-Template Matching sebagai pendeteksian objek di depan yang diintegrasikan dengan perhitungan Stereoscopic Measurement mendapatkan skor Mean Absolute Percentage Error (MAPE) sebesar 6,009% dalam deteksi dan estimasi jarak antara perahu.

Kata Kunci: *Autonomous Ship, ZED Stereo Camera, YOLOv5, Template Matching, Stereoscopic Measurement*

ABSTRACT

DANDY GARDA DIRGANTARA. *Relative Distance Detection System On Autonomous Ship With Stereoscopic Measurement* (supervised by Prof. Dr. Indrabayu, ST., MT., M.Bus.Sys., IPM., ASEAN Eng. and Dr.Eng. Faisal Mahmuddin, ST., M.Inf.Tech., M.Eng)

Autonomous Ship technology or unmanned ships is a technology that can operate ships without requiring input from humans, by working with artificial intelligence (Artificial Intelligence) to see the situation around the ship. In the development of an Autonomous Ship, the obstacle that must be faced is how close the ship is to the object in front of it, for that a system is needed that can detect the relative distance in real-time. In this research, the Computer Vision method will be used which can provide information regarding surrounding objects such as fishing boats and is combined in Stereoscopic Measurement calculations with a stereo camera, namely the ZED Stereo Camera, so that the distance between boats can be known directly. And in this study, the detection of the distance between the boats from 7 to 20 meters was carried out. The results of the study show that the use of the YOLOv5-Template Matching method as the detection of objects in front which is integrated with the Stereoscopic Measurement calculations obtains a Mean Absolute Percentage Error (MAPE) score of 6.009% in the detection and estimation of the distance between boats.

Keywords: Autonomous Ship, ZED Stereo Camera, YOLOv5, Template Matching, Stereoscopic Measurement

DAFTAR ISI

LEMBAR PENGESAHAN SKRIPSI	i
PERNYATAAN KEASLIAN.....	ii
ABSTRAK.....	iii
ABSTRACT.....	iv
DAFTAR ISI.....	v
DAFTAR GAMBAR	vii
DAFTAR TABEL.....	viii
DAFTAR SINGKATAN DAN ARTI SIMBOL	ix
DAFTAR LAMPIRAN.....	x
KATA PENGANTAR	xi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan Penelitian.....	2
1.4 Manfaat Penelitian.....	2
1.5 Ruang Lingkup	3
1.6 Sistematika Penulisan.....	3
BAB II TINJAUAN PUSTAKA.....	5
2.1 Autonomous Ship.....	5
2.2 Citra	5
2.3 Tipe Citra.....	6
2.3.1 Citra Biner.....	6
2.3.2 Citra Grayscale.....	7
2.3.3 Citra Warna	7
2.4 Resolusi Citra	8
2.5 Computer Vision	9
2.6 Deteksi Objek.....	9
2.7 Python.....	10
2.8 OpenCV.....	11
2.9 ZED Stereo Camera.....	12
2.10 PyTorch.....	12
2.11 Jaringan Syaraf Tiruan.....	13
2.12 Activation Function	15
2.13 Algoritma You Only Look Once	17

2.14	Template Matching	22
2.15	Stereoscopic Measurement	22
2.16	Penelitian Terkait	23
BAB III METODE PENELITIAN.....		26
3.1	Tahapan Penelitian	26
3.2	Waktu dan Lokasi Penelitian.....	27
3.3	Instrumen Penelitian	27
3.4	Teknik Pengambilan Data	28
3.5	Perancangan Implementasi Sistem.....	29
3.5.1	Input Citra	31
3.5.2	Preprocessing	33
3.5.3	Deteksi Objek Menggunakan YOLOv5.....	34
3.5.4	Deteksi Objek Menggunakan <i>Template Matching</i>	38
3.5.5	Konfigurasi YAML.....	38
3.5.6	Stereoscopic Measurement.....	38
3.5.7	Evaluasi Sistem	42
BAB IV HASIL DAN PEMBAHASAN		44
4.1	Hasil dan Pembahasan Model YOLOv5	44
4.2	Hasil dan Pembahasan Sistem Deteksi Jarak	47
BAB V KESIMPULAN & SARAN		51
5.1	Kesimpulan.....	51
5.2	Saran	51
DAFTAR PUSTAKA		53
LAMPIRAN.....		56

DAFTAR GAMBAR

Gambar 2. 1 Citra Biner dengan nilai 0 dan 1.....	7
Gambar 2. 2 Citra Grayscale dengan nilai piksel antara 0-255.....	7
Gambar 2. 3 Citra Gambar dengan nilai RGB	8
Gambar 2. 4 ZED Stereo Camera	12
Gambar 2. 5 Arsitektur Neural Network (Harani & Hasanah, 2020)	13
Gambar 2. 6 Linear Function	16
Gambar 2. 7 Fungsi Aktivasi ReLU (Sharma, 2017).....	16
Gambar 2. 8 Fungsi Aktivasi Leaky ReLU (Arif, 2022)	17
Gambar 2. 9 Arsitektur YOLOv5 (Nguyen dkk., 2021)	19
Gambar 2. 10 Struktur Focus pada YOLOv5 (Jing dkk., 2022)	20
Gambar 2. 11 Struktur Module BottleneckCSP.....	20
Gambar 2. 12 Struktur Module Bottleneck	21
Gambar 2. 13 Struktur Module SPP.....	21
Gambar 2. 14 Proses dari Template Matching.....	22
Gambar 2. 15 Ilustrasi teknik stereoscopic	23
Gambar 3. 1 Tahapan Penelitian	26
Gambar 3. 2 Contoh perahu nelayan sebagai data training.....	28
Gambar 3. 3 Ilustrasi Pengambilan Testing	29
Gambar 3. 4 Rancangan Sistem Deteksi Jarak Pada Perahu.....	30
Gambar 3. 5 ZED Stereo Camera	31
Gambar 3. 6 Penampakan pada Stereo Camera	31
Gambar 3. 7 Hasil Konversi Warna RGB ke Grayscale	34
Gambar 3. 8 Flowchart YOLOv5	37
Gambar 3. 9 Pengambilan jarak objek dengan Stereoscopic	39
Gambar 3. 10 Posisi objek dan kamera yang berbeda	40
Gambar 4. 1 Grafik nilai mAP pada setiap epoch.....	44
Gambar 4. 2 Grafik nilai precision pada setiap epoch	45
Gambar 4. 3 Grafik nilai recall pada setiap epoch.....	46
Gambar 4. 4 Deteksi jarak pada perahu nelayan.....	47
Gambar 4. 5 Persentase Error Jarak Prediksi	49
Gambar 4. 6 Contoh hasil kesalahan pembentukan bounding box	50
Gambar 4. 7 Hasil 2 frame yang digabungkan.....	50

DAFTAR TABEL

Tabel 2. 1 Penelitian Terkait	24
Tabel 3. 1 Range Perhitungan MAPE	43
Tabel 4. 1 Evaluasi sistem deteksi jarak pada perahu	48

DAFTAR SINGKATAN DAN ARTI SIMBOL

Lambang/Singkatan	Arti dan Keterangan
AV	<i>Autonomous Vehicle</i>
IoU	<i>Intersection over Union</i>
MAPE	<i>Mean Absolut Percentage Error</i>
MSC	<i>Maritime Safety Committee</i>
RGB	<i>Red, Green, Blue</i>
RMSE	<i>Root Mean Squared Error</i>
USV	<i>Unmanned Surface Vehicle</i>
YOLO	<i>You Only Look Once</i>

DAFTAR LAMPIRAN

Lampiran 1. Beberapa Contoh Dataset Primer	56
Lampiran 2. Hasil Deteksi Jarak Dalam 100 Frame	58
Lampiran 3. Source Code.....	63
Lampiran 4. Hasil Training di Google Colab.....	78

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa yang telah memberikan rahmat, taufik dan hidayah-Nya, sehingga penulis dapat menyelesaikan laporan tugas akhir dengan berjudul *“Sistem Deteksi Jarak Relatif Pada Autonomous Ship Dengan Stereoscopic Measurement”*. Laporan Tugas Akhir ini merupakan salah satu syarat untuk dapat memperoleh gelar Program Sarjana Strata Satu (S1) di Departemen Teknik Informatika Fakultas Teknik Universitas Hasanuddin Makassar.

Penulis menyadari bahwa dalam penyusunan Laporan Tugas Akhir ini, penulis mendapatkan banyak sumbangan pikiran serta bimbingan, baik secara moral maupun materi dari beberapa pihak. Sehingga pada kesempatan ini, penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

- Kedua orang tua penulis, Bapak Ir. Misbahuddin Hamid dan Ibu Yuyu Srie Rahayu, SH., kakak tercinta Bayu Aditya Nugraha yang telah memberikan dukungan, doa dan semangat yang tiada hentinya, serta selalu sabar dalam mendidik penulis sejak kecil.
- Bapak Dr. Indrabayu, ST., MT., M.Bus.Sys., IPM. sebagai pembimbing I dan Bapak Dr.Eng. Faisal Mahmuddin, ST., M.Inf.Tech., M.Eng. sebagai pembimbing II yang senantiasa menyediakan waktu, tenaga, pikiran dan perhatian yang luar biasa dalam mengarahkan penulis untuk menyelesaikan tugas akhir.
- Kakak-kakak dari AIMP Research Group Unhas yaitu Kak Muhammad Fadhil Bin Bahrunnida, Kak Nublan Azqalani, Kak Herul Hidayat, Kak Taslinda, Kak Muhammad Abdillah Rahmat dan Kak Irfan Ripat yang telah memberikan begitu banyak bantuan selama penelitian, pengambilan data hingga diskusi progress penyusunan Tugas Akhir.
- Muhammad Fandly Fadlurachman, Andi Andhika Pangerang Pallampa, Fadilah Ramadhani, Maghfirah Tenri Sumpala Zani, Nur Hasana Abunawas dan Jabalnur, menjadi kawan yang memiliki tujuan yang sama dan saling menguatkan agar dapat menyelesaikan Tugas Akhir.

- Darul Ikhsan, terima kasih telah membantu penulis dalam pengambilan data testing di atas perahu.
- Sahabat penulis, Budi Ginanjar dari Bilalang 3 Kotamobagu. Terima kasih untuk tetap menjadi teman yang menyenangkan dan penuh tawa.
- Mas Airlangga Rahmadi dan Mas Ilham Rahmadi, terima kasih atas ilmu serta menjadi teman ngobrol penulis selama pengerjaan Tugas Akhir ini.
- Segenap Dosen dan Staf Departemen Teknik Informatika Fakultas Teknik Universitas Hasanuddin yang telah membantu dalam kelancaran penyelesaian Tugas Akhir.
- Teman-teman seangkatan Synchronous, terima kasih atas dukungan dan semangat yang diberikan selama ini.
- Serta pihak-pihak lain yang tidak disebutkan dan tanpa sadar telah menjadi inspirasi penulis.

Akhir kata, penulis berharap semoga Tuhan Yang Maha Esa berkenan membalas segala kebaikan dari semua pihak yang telah banyak membantu. penulis menyadari masih terdapat kekurangan dalam penyusunan Laporan Tugas Akhir ini baik isi maupun cara penyajian. Oleh dikarenakan itu penulis mengharapkan adanya bentuk saran serta masukan yang membangun demi kesempurnaan laporan ini. Penyusun berharap semoga Laporan Tugas Akhir ini dapat memberikan manfaat bagi pembaca pada umumnya dan manfaat bagi penulis pada khususnya. Aamiin.

Gowa, Februari 2023

Penulis,
Dandy Garda Dirgantara

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kendaraan tanpa awak (*Autonomous Vehicle – AV*) merupakan salah satu moda transportasi kendaraan cerdas yang menggabungkan ilmu teknik komputer yang terintegrasi dan memiliki fitur otomatisasi untuk beroperasi tanpa memerlukan input dari manusia (Arifianto dkk., 2021). Teknologi AV menggunakan kecerdasan buatan (*Artificial Intelligence*) sehingga dapat melihat objek sekelilingnya (Ardi & Susilowati, 2022), Salah satu dari pengembangan *Autonomous Vehicle* adalah kapal tanpa awak atau *Autonomous Ship*.

Sebuah teknologi *Autonomous Ship* dapat digunakan untuk berbagai tugas di area aplikasi yang berbeda seperti pengawasan intelijen pantai, keamanan pelabuhan dan perbatasan, transmisi sinyal antara kendaraan udara dan bawah air (Zhu & Wen, 2019).

Menurut *Maritime Safety Committee* (MSC), dalam pengembangan *Autonomous Ship* terdapat 4 tahap status pengembangan, dimana untuk mencapai tahap 1 yaitu “Kapal dengan proses otomatis dan dukungan keputusan”, diperlukan sistem yang dapat mendeteksi jarak relatif antar kapal (Chae dkk., 2020).

Dalam mewujudkan pendeteksian jarak relatif, hal pertama yang diperlukan adalah pendeteksian objek yang di depan dengan jelas dan cepat untuk menghindari rintangan, mengikuti jalur atau target, membaca rambu atau berinteraksi dengan kapal lain (Lee dkk., 2018). Oleh karena itu, untuk mengetahui seberapa dekat rintangan atau objek dengan kapal, diperlukan sebuah sistem yang dapat mendeteksi jarak relatif secara *real-time*.

Penghitungan jarak relatif antara objek dapat dilakukan dengan pengambilan gambar stereoskopik (*stereo camera*) dan mengambil nilai-nilai dari spesifikasi kamera yang dapat digunakan dalam menentukan jarak relatif antara objek (posisi *stereo camera*) dan objek yang ada didepannya (Dawood dkk., 2017).

Stereo camera bekerja dengan mengikuti mata manusia dan juga dapat menjangkau bagian kiri dan kanan perahu sehingga pengambilannya lebih luas, serta menciptakan gambar 3D dengan menggabungkan 2 gambar dari 2 lensa yang

berbeda (Dawood dkk., 2017). Oleh karena itu, akan menggunakan ZED Stereo Camera yang dapat melakukan 2 hal sekaligus yaitu mendeteksi objek di depan beserta pemberian nama objek/*class* dan penghitungan jarak relatif antara kapal pengamat dan objek.

Pada tahap pengembangan awal untuk *Autonomous Ship*, hal yang dilakukan dalam penelitian ini dimulai dengan pendeteksian perahu terlebih dahulu dengan menggunakan teknik deep learning yaitu YOLOv5 untuk mempelajari citra dan mengidentifikasi objek di dalam gambar, dimana YOLOv5 mampu melakukan deteksi objek hanya dengan melihat gambar sekali, sehingga dapat bekerja dengan cepat dan efisien. Selanjutnya hasil dari YOLOv5 dapat sebagai acuan untuk mendapatkan jarak relatif antara *stereo camera* dan objek dengan pengukuran Stereoscopic Measurement.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang dijelaskan, maka rumusan masalah pada tugas akhir ini adalah:

- a. Bagaimana cara mengembangkan sistem deteksi jarak relatif pada *autonomous ship* dengan *stereo camera*?
- b. Bagaimana hasil kinerja sistem deteksi jarak relatif jika ditempatkan di *autonomous ship* dengan *stereo camera*?

1.3 Tujuan Penelitian

Untuk bagian tujuan masalah, penulis membagi menjadi beberapa poin:

- a. Untuk mengembangkan sistem deteksi jarak relatif pada *autonomous ship* dengan *stereo camera*
- b. Untuk mengetahui performa sistem deteksi jarak relatif pada *autonomous ship* dengan *stereo camera*.

1.4 Manfaat Penelitian

Untuk manfaat yang dapat diberikan dari penelitian ini adalah:

- a. Bagi pelaut, sistem ini dapat digunakan untuk mendeteksi jarak antara perahu sehingga dapat membantu menghindari tabrakan.

- b. Bagi peneliti dapat digunakan sebagai dasar dalam melakukan riset berkelanjutan mengenai industri *autonomous ship*.

1.5 Ruang Lingkup

Untuk bagian ruang lingkup pada penelitian ini, penulis membagi menjadi beberapa poin:

- a. Data training yang digunakan adalah data primer dari beberapa tempat di wilayah perairan area Makassar.
- b. Testing dilakukan pada siang hari di Sungai Jeneberang, Jembatan Barombong, Kec. Tamalate, Kota Makassar
- c. Perahu pengamat digunakan untuk menempatkan *stereo camera* dan laptop.
- d. 1 objek perahu nelayan (*fishing boat*) berada di depan dengan berbagai posisi relatif dari perahu pengamat.
- e. Jarak perahu pengamat dan perahu nelayan (*fishing boat*) bervariasi dari 7-20 meter, dengan perahu yang diamati berada di depan.
- f. Evaluasi keakuratan sistem deteksi jarak menggunakan metode MAPE (*Mean Absolute Percentage Error*)

1.6 Sistematika Penulisan

Pada bagian ini memberikan gambaran singkat mengenai isi tulisan secara keseluruhan pada tugas akhir ini, maka akan diuraikan beberapa tahapan dari penulisan secara sistematis, yaitu:

BAB 1 PENDAHULUAN

Dalam bab ini akan menjelaskan tentang latar belakang diangkatnya judul penelitian Sistem Deteksi Jarak Relatif pada Autonomous Ship dengan Stereoscopic Measurement, disertai dengan rumusan masalah, tujuan penelitian, manfaat penelitian, ruang lingkup dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Dalam bab ini berisi teori-teori yang berhubungan dengan proses penelitian yang dilakukan.

BAB III METODOLOGI PENELITIAN

Dalam bab ini berisi tentang perencanaan dan penerapan algoritma serta teknik pengolahan data.

BAB IV HASIL DAN PEMBAHASAN

Dalam bab ini berisi tentang hasil pengolahan data serta pembahasan yang disertai tabel hasil penelitian.

BAB V PENUTUP

Dalam bab ini berisi tentang kesimpulan yang didapatkan berdasarkan hasil penelitian yang dilakukan, serta memberikan saran-saran untuk penelitian/pengembangan selanjutnya.

BAB II

TINJAUAN PUSTAKA

2.1 Autonomous Ship

Teknologi *autonomous* secara bertahap akan membentuk kembali industri maritim dan mempunyai potensi besar dalam memberikan manfaat. Contohnya keunggulan dalam teknologi ini adalah penghematan bahan bakar, peningkatan efisiensi dan produktivitas pada biaya, dan dapat mengoptimalkan rantai logistik melalui solusi teknologi informasi (TI) (Hagen, 2021).

Unmanned Surface Vehicle (USV) atau Kapal tanpa awak adalah bagian dari *Autonomous Ship* yang merupakan kendaraan tanpa awak/kru yang dapat dioperasikan di atas permukaan air. USV umumnya selain digunakan sebagai kapal riset juga dapat digunakan sebagai kapal survai, inspeksi keadaan perbatasan laut dan juga bisa digunakan sebagai alat operasi penyelamatan (Fikri & Rivai, 2020).

Dalam pengembangan *Autonomous Ship*, ada 4 tahap status yang diperlukan menurut *Maritime Safety Committee (MSC)*: Pertama, kapal dengan proses otomatis dan dukungan keputusan. Kedua, kapal yang dikendalikan dari jarak jauh dengan awak kapal di dalamnya. Ketiga, kapal yang dikendalikan dari jarak jauh tanpa awak kapal di dalamnya. Dan terakhir, kapal sudah menjadi sepenuhnya *autonomous* (Chae dkk., 2020).

2.2 Citra

Secara umum, citra merupakan suatu gambar, foto ataupun berbagai tampilan dua dimensi yang menggambarkan visualisasi objek. Citra dapat dikategorikan ke dalam citra kontinu dan citra digital. Citra kontinu dihasilkan dari sistem optik yang menerima sinyal analog, contohnya adalah mata manusia, kamera analog. Citra digital dihasilkan melalui proses digitalisasi terhadap citra kontinu (Zonyfar, 2020).

Citra digital merupakan representasi dari fungsi intensitas cahaya dalam bentuk diskrit pada bidang dua dimensi. Citra digital tersimpan dalam bentuk larik (*array*) angka digital yang merupakan hasil kuantifikasi dari tingkat kecerahan

masing-masing piksel. Struktur *array* ini tersusun dalam baris horisontal dan kolom vertikal (*samples*) (Prabowo & Abdullah, 2018).

Citra digital dapat diwakili oleh sebuah matriks yang terdiri dari beberapa kolom dan beberapa baris, dimana perpotongan antara kolom dan baris tersebut disebut dengan *pixel* (*picture element*) yang merupakan elemen terkecil dari sebuah citra. Secara matematis citra digital dapat dituliskan sebagai intensitas $f(x,y)$ dimana nilai x (baris) dan y (kolom) adalah koordinat posisi dan $f(x,y)$ merupakan nilai fungsi pada setiap (x,y) menyatakan besar intensitas yang diterima oleh citra atau tingkat keabuan atau warna dari piksel di titik tersebut (Munantri dkk., 2020).

2.3 Tipe Citra

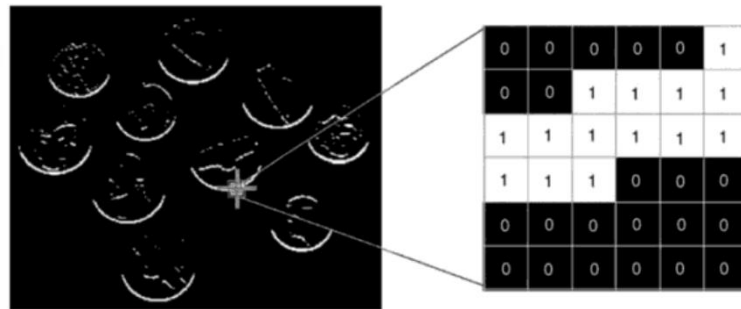
Dalam citra terdapat beberapa tipe yang digunakan dalam penelitian, diantaranya adalah:

2.3.1 Citra Biner

Tiap-tiap piksel dalam citra biner hanya membutuhkan 1 bit memori. Maka, setiap piksel hanya mempunyai dua buah nilai intensitas yaitu 1 atau 0 (Andono dkk., 2017).

Nilai intensitas 0 berarti menunjukkan warna hitam dan 1 berarti menunjukkan putih. Keuntungan utama dari representasi ini adalah citra biner sangatlah cocok untuk citra yang memuat grafika sederhana, teks atau garis dengan ukurannya yang kecil.

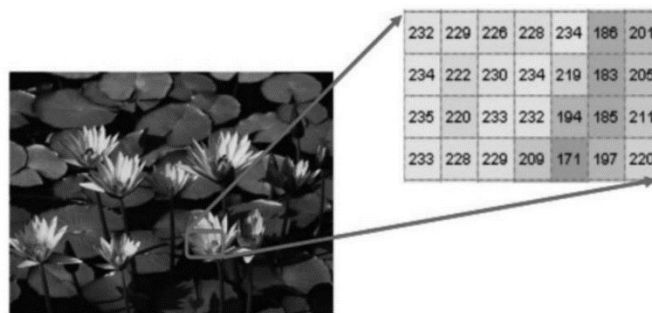
Dalam Gambar 2.1 menunjukkan sebuah citra biner dan sebuah region berukuran 6×6 , dimana piksel dengan nilai 1 berkaitan dengan tepi-tepi dan piksel dengan nilai 0 berkaitan dengan latar gambar (Siahaan & Sianipar, 2020).



Gambar 2. 1 Citra Biner dengan nilai 0 dan 1

2.3.2 Citra Grayscale

Citra grayscale merupakan matriks data yang nilai-nilainya mewakili intensitas setiap piksel berkisar antara 0 sampai 255 dan dalam piksel membutuhkan 8 bit memori, seperti pada Gambar 2.2 di bawah ini (Andono dkk., 2017).

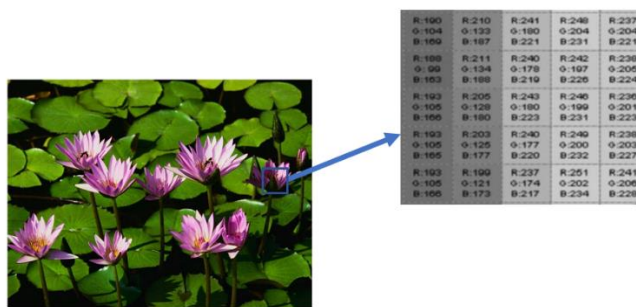


Gambar 2. 2 Citra Grayscale dengan nilai piksel antara 0-255

Warna yang dimiliki dari citra grayscale adalah hitam, keabuan dan putih. Tingkatan keabuan di sini merupakan warna abu dengan berbagai tingkatan dari hitam hingga mendekati putih (Putra, 2010).

2.3.3 Citra Warna

Dalam citra warna memiliki 3 komponen di setiap masing-masing piksel yaitu komponen merah (*red*), hijau (*green*) dan biru (*biru*). Tiap intensitas warna merah, hijau dan biru dikombinasikan dan disimpan pada lokasi piksel, seperti contoh pada Gambar 2.3 di bawah ini. (Andono dkk., 2017).



Gambar 2. 3 Citra Gambar dengan nilai RGB

Dengan citra warna 24 bit (RGB) menggunakan tiga buah array 2D yang berukuran sama, setiap elemen array memuat nilai 8-bit, yang mengindikasikan intensitas dari merah, hijau atau biru dengan skala [0, 255]. Jika dikombinasikan ketiga nilai itu maka menghasilkan nilai 24 bit atau 16 juta kombinasi warna (Siahaan & Sianipar, 2020).

2.4 Resolusi Citra

Resolusi citra merupakan tingkat detail dalam suatu citra. Semakin tingginya resolusi citra maka akan semakin tinggi tingkat detail dari citra tersebut. Satuan yang digunakan dalam resolusi citra dapat berupa ukuran fisik yaitu jumlah garis per mm/jumlah garis per inci atau dapat berupa ukuran citra menyeluruh (jumlah garis per tinggi citra) (Putra, 2010).

Resolusi piksel merupakan resolusi yang sering digunakan dalam citra grafis atau gambar. Semakin tinggi suatu resolusi, maka makin halus (tajam) rincian yang bisa dilihat. Tampilan grafis dengan ukuran besar dan memiliki kualitas yang baik akan memerlukan resolusi yang tinggi (Azis, 2011).

Dalam resolusi piksel, sebuah citra dengan tinggi N piksel dan lebar M piksel berarti memiliki resolusi sebesar M x N. Resolusi piksel akan memberikan dua buah angka integer yang secara berurutan akan mewakili jumlah piksel lebar dan jumlah piksel tinggi dari citra tersebut. Jenis resolusi ini sering kali dijumpai dalam kamera digital, suatu citra yang memiliki lebar 2048 piksel dan tinggi 1536 piksel maka akan memiliki total piksel sebanyak $2048 \times 1536 = 3.145.728$ piksel (Putra, 2010).

2.5 Computer Vision

Computer vision adalah sebuah ilmu pengetahuan dalam bidang ilmu informatika yang menggunakan komputer sebagai media untuk dapat melihat objek dari jarak terdekat. Dalam *computer vision* dapat digunakan sebagai media pembelajaran, pengenalan suara, pengenalan wajah dan digunakan diberbagai bidang seperti industri pendidikan, kesehatan serta pertahanan dan keamanan (Lubis dkk., 2020).

Hal yang sangat berkaitan erat dengan pengolahan citra adalah *computer vision*. Pada hakikatnya, *computer vision* mencoba meniru cara kerja sistem visual manusia (*human vision*). *Human vision* sangatlah kompleks, manusia melihat objek dengan indera penglihatan (mata), lalu citra objek diteruskan ke otak untuk diinterpretasi sehingga manusia mengerti objek apa yang tampak dalam pandangan matanya. Hasil interpretasi ini mungkin digunakan untuk pengambilan keputusan.

Computer vision adalah proses otomatis yang mengintegrasikan sejumlah besar proses untuk persepsi visual, seperti akuisisi citra, pengolahan citra (*preprocessing*), klasifikasi, pengenalan pola (*recognition*), dan membuat keputusan. *Computer vision* terdiri dari teknik-teknik untuk mengestimasi ciri-ciri objek di dalam citra, pengukuran ciri yang berkaitan dengan geometri objek, dan menginterpretasi informasi geometri tersebut.

Proses-proses yang dilakukan dalam *computer vision* dapat dibagi menjadi tiga aktivitas (Widodo, 2021):

1. Memperoleh atau mengakuisisi citra digital.
2. Melakukan teknik komputasi untuk memproses atau memodifikasi data citra dengan menggunakan operasi-operasi pengolahan citra
3. Menganalisis dan menginterpretasi citra dan menggunakan hasil pemrosesan untuk tujuan tertentu, misalnya memandu robot, mengontrol peralatan, memantau proses manufaktur, dan lain-lain

2.6 Deteksi Objek

Deteksi objek (*Object Detection*) merupakan teknik visi komputer untuk menemukan contoh objek dalam gambar atau video. Algoritma deteksi objek pada umumnya memanfaatkan pembelajaran mesin (*Machine Learning*) dan

pembelajaran mendalam (*Deep Learning*) untuk menghasilkan hasil yang bermakna. Saat manusia melihat gambar atau video, manusia dapat mengenali dan menemukan objek dalam beberapa saat berbeda dengan komputer yang memerlukan komputasi yang kompleks.

Tujuan deteksi objek adalah untuk mereplikasi kecerdasan yang dimiliki manusia dalam melihat benda menggunakan komputer. Cara kerja deteksi objek adalah deteksi objek menempatkan keberadaan objek dalam gambar dan menggambar kontak pembatas di sekitar objek itu. Biasanya melibatkan dua proses yaitu mengklasifikasi jenis objek kemudian menggambar kotak di sekitar objek itu. Secara umum, klasifikasi adalah mengklasifikasikan gambar ke dalam kategori tertentu. Sedangkan objek deteksi adalah mengidentifikasi lokasi objek dalam gambar (Aningtiyas dkk., 2020).

Metode untuk deteksi objek umumnya jatuh ke dalam pendekatan berbasis pembelajaran mesin atau pendekatan berbasis pembelajaran dalam. Dalam teknik pembelajaran yang mendalam mampu melakukan deteksi objek ujung ke ujung secara khusus mendefinisikan fitur dan biasanya didasarkan pada jaringan saraf convolutional (CNN).

Pertimbangan utama yang harus diperhatikan saat memilih antara *Machine Learning* dan *Deep Learning* adalah apakah memiliki GPU yang kuat dan banyak gambar pelatihan berlabel. Dalam teknik Deep Learning cenderung bekerja lebih baik ketika memiliki lebih banyak gambar, dan GPU mengurangi waktu yang dibutuhkan untuk melatih model (Putro dkk., 2020).

Untuk melatih detektor objek, diperlukan merancang arsitektur jaringan untuk mempelajari fitur untuk objek yang menarik. Diperlukan kumpulan data berlabel yang sangat besar untuk melatih CNN dan diperlukan secara manual mengatur layer dan bobot di CNN yang membutuhkan banyak waktu dan data penelitian (Putro dkk., 2020).

2.7 Python

Python merupakan bahasa pemrograman tingkat tinggi yang sangat populer saat ini, dikarenakan tidak terlepas dari bahasa python yang dianggap *powerful* dan

mendekati bahasa manusia. Python sangat dibutuhkan di bidang *Data Science*, *Machine Learning*, *Deep Learning*, *Face Recognition* dan bidang lainnya.

Python juga mengikuti bahasa pemrograman “*interpreter*”, yang artinya kode akan langsung dieksekusi sesuai instruksi yang ditulis dalam bahasa pemrograman atau *scripting* tanpa sebelumnya mengubahnya menjadi kode objek seperti compiler (Rangkuti dkk., 2021).

Python memiliki banyak *library* (modul-modul siap pakai) yang mendukung kegiatan *Machine Learning* atau *Deep Learning* yaitu **Matplotlib**, **Numpy**, **Pandas**, **Scikit-learn** dan *library* lainnya. Dahulu, program semacam itu harus ditulis ulang di bahasa lain seperti Java atau C++ agar dapat bekerja dengan efisien, sekarang dengan Python dapat digunakan baik untuk keperluan eksperimental maupun langsung diterapkan di lingkungan produksi (Kurniawan, 2022).

Bahasa tingkat tinggi memiliki kelebihan yaitu mudah dipelajari, mudah ditulis, mudah dibaca dan juga mudah dicari kesalahan dalam kode. Namun, hal yang menjadi kekurangannya adalah memerlukan waktu untuk memproses suatu program sebelum program tersebut dijalankan (Utami dkk., 2004).

2.8 OpenCV

OpenCV merupakan singkatan dari *Open Source Computer Vision* adalah sebuah *library* yang dikhususkan untuk melakukan penglihatan komputer secara real time/waktu sebenarnya. Dikembangkan oleh pusat penelitian Intel di Nizhny Novgorod, Rusia. Kata Open pada library OpenCV adalah open source yakni gratis, tidak berbayar dan dapat didownload oleh siapa saja. Sedangkan untuk kata CV merupakan singkatan dari *Computer Vision*, dimana komputer yang digunakan untuk mengolah/mengerjakan image atau citra/gambar yang ditangkap oleh alat perekam seperti webcam atau kamera yang dikonversi dari bentuk analog ke dalam bentuk digital lalu dioleh dalam komputer tersebut. Tujuan dari *library* ini yaitu dapat digunakan untuk memperbaiki/menyempurnakan dari kualitas gambar atau untuk dilakukan proses identifikasi pada gambar.

Dalam fitur utama OpenCV, *library* ini dapat digunakan untuk pengolahan citra digital (untuk *low* dan *mid-level API*) dimana melakukan eksperimen uji coba

dengan berbagai standar algoritme yaitu deteksi pada garis, tepi, pucuk, proyeksi elips, image pyramid untuk melakukan pemrosesan gambar multi skala, pencocokkan template, dan juga berbagai transform (Fourier, cosine diskrit, distance transform) dan lain-lain, pengolahan *high-level* yaitu keahlian dalam melakukan deteksi pada wajah, pengenalan wajah dan termasuk juga *optical flow*, serta digunakan dalam *machine learning* dan transformasi gambar (Harani & Hasanah, 2020).

2.9 ZED Stereo Camera

ZED merupakan kamera yang mereproduksi cara kerja penglihatan manusia. Menggunakan dua “mata” dan melalui triangulasi, ZED memberikan pemandangan tiga dimensi yang diamatinya, memungkinkan sadar ruang dan gerak.



Gambar 2. 4 ZED Stereo Camera

Seperti yang terlihat pada Gambar 2.4. ZED juga memiliki lensa ganda dimana dapat menangkap video 3D dalam definisi tinggi dengan bidang pandang yang luas dan menghasilkan dua aliran video kiri dan kanan yang disinkronkan dalam format berdampingan pada USB 3.0 (Stereolabs, 2020).

2.10 PyTorch

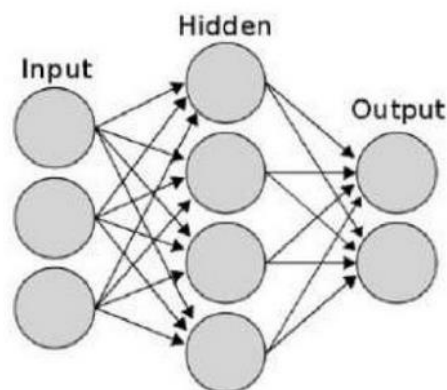
PyTorch merupakan *library machine learning/deep learning* berbasis sumber terbuka untuk Python yang berfokus dalam pembelajaran mesin. *Library* ini dikembangkan oleh kelompok riset AI (*Artificial Intelligence*) dari Facebook. PyTorch ini memungkinkan/mengharuskan setiap pengembang/developer untuk melakukan proses perhitungan *tensor* dengan menggunakan akselerasi GPU,

membuat grafik komputasi dinamis dan juga dapat menghitung gradien secara otomatis (Harani & Hasanah, 2020).

Dalam PyTorch menyediakan struktur data inti yang disebut dengan Tensor, sebuah array multidimensional yang memiliki fitur kemiripan dengan library NumPy. Banyak fitur yang dapat dibangun untuk mempermudah pembuatan proyek dan menjalankannya atau untuk mendesign dan melatih percobaan terhadap arsitektur baru *Neural Network* (Jaringan Saraf Tiruan). Tensor pun dapat digunakan mengakselerasi operasi matematika (asumsi *hardware* dan *software* yang digunakan merupakan teknologi baru). *Library* ini pun dapat melakukan training secara terpisah (*distributed training*), *worker process* untuk menggunakan beberapa sumber daya pada satu mesin sebagai efisiensi pemuatan data, dan *library* tambahan dari fungsi *Deep Learning* pada umumnya (Hoki dkk., 2021)

2.11 Jaringan Syaraf Tiruan

Jaringan Syaraf Tiruan/*Neural Network* adalah teknik menirukan syaraf manusia yang merupakan bagian fundamental dari otak. *Neural Network* terdiri atas lapis masukan (*input layer*) dan lapis keluaran (*output layer*). Setiap lapis terdiri atas satu atau beberapa unit neuron yang mempunyai sebuah fungsi aktivasi yang menentukan keluaran dari unit tersebut. Dalam *neural network* dapat menambahkan lapis tersembunyi (*hidden layer*) untuk menambah kemampuan dari NN tersebut (Batubara & Awangga, 2020). Arsitektur *Neural Network* dapat diilustrasikan sebagai Gambar 2.5 di bawah berikut ini:



Gambar 2. 5 Arsitektur Neural Network (Harani & Hasanah, 2020)

Pada Gambar 2.5 menunjukkan arsitektur dari *Neural Network* yang merupakan gambaran susunan komponen *layer* dan *neuron* pada *input*, *hidden* dan *output* yang terhubung dengan bobot, fungsi aktivasi dan fungsi pembelajaran (*learning function*). Berikut penjelasan pada masing-masing lapisan (Harani & Hasanah, 2020):

1. Lapisan Masukan (*Input Layer*)

Lapisan masukan merupakan lapisan yang terdiri dari sejumlah neuron yang akan menerima/menampung sinyal dari luar dan kemudian akan menyampaikan ke neuron-neuron lainnya dalam jaringan.

2. Lapisan Tersembunyi (*Hidden Layer*)

Lapisan tersembunyi berguna untuk meningkatkan kemampuan/kapasitas jaringan dalam menyelesaikan suatu masalah. Dampak dari lapisan ini adalah pelatihan (*training*) menjadi makin sulit atau lama.

3. Lapisan Keluaran (*Output Layer*)

Lapisan keluaran berguna untuk menyampaikan sinyal-sinyal keluaran atau *output* hasil pemrosesan jaringan.

Neural Network dapat dilatih dengan menggunakan data *training*. Dengan banyaknya data training maka akan semakin bagus hasil kerja dari NN tersebut. Namun, karena kemampuan dari NN juga terbatas pada jumlah lapisan, dengan semakin banyak jumlah lapisan maka semakin tinggi kapasitas NN tersebut. Banyaknya lapisan juga membawa kekurangan yaitu semakin banyaknya jumlah iterasi atau training yang dibutuhkan. Untuk dapat mengatasi kekurangan ini, dikembangkanlah teknik *Deep Learning*.

Neural Network bekerja dengan mengambil prinsip dari kemampuan otak manusia yang mampu memberikan stimulasi/rangsangan, melakukan sebuah proses dan memberikan *output*. *Output* yang diperoleh dari variasi rangsangan dan proses yang terjadi di dalam otak manusia. Kemampuan manusia dalam memproses sebuah informasi adalah hasil kompleksitas proses di dalam otak. Yang bisa dicontohkan dalam kehidupan nyata adalah anak-anak, mereka mampu belajar untuk melakukan pengenalan sesuatu meskipun mereka tidak mengetahui algoritma apa yang mereka gunakan untuk pengenalan. Dikarenakan kekuatan komputasi

yang luar biasa dari otak manusia ini merupakan sebuah keunggulan di dalam ilmu pengetahuan (Batubara & Awangga, 2020).

Fungsi yang dilakukan dalam *Neural Network* adalah:

- Pengklasifikasian pola.
- Mengatur pola yang didapat dari *input* ke dalam pola baru pada *output*.
- Penyimpan pola yang akan dipanggil kembali.
- Memetakan pola-pola yang sejenis.
- Pengoptimasi permasalahan.
- Prediksi.

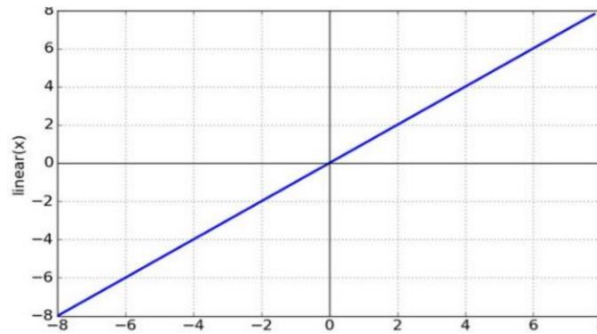
2.12 Activation Function

Dalam *neural network*, fungsi aktivasi (*activation function*) berfungsi untuk menentukan apakah neuron tersebut harus “aktif” atau tidak, dengan berdasarkan *weighted sum* dari inputan. Pada umumnya, fungsi aktivasi memiliki 2 jenis dalam *neural network*, yaitu *Linear* dan *Non-linear activation function* (Batubara & Awangga, 2020):

1. Linear Function

Fungsi aktivasi dalam bentuk *default* adalah Linear. Jika neuron menggunakan *linear function*, maka keluaran dari neuron tersebut adalah *weighted sum* dari input dan bias. Grafik *linear function* dapat dilihat pada Gambar 2.6.

Pada Gambar 2.6, *linear function* memiliki persamaan $f(x) = x$ dengan rentang (-infinity to infinity) yang berarti fungsi ini tidak membantu dalam kompleksitas atau berbagai parameter data biasa yang diberikan ke *neural network* (Sharma, 2017).



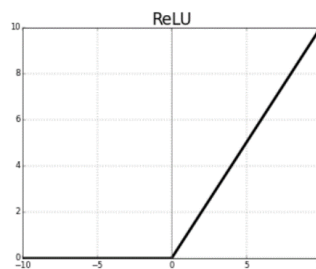
Gambar 2. 6 Linear Function

2. *Non-linear Function*

Fungsi *Non-linear function* merupakan *activation function* yang sering digunakan, fungsi ini dapat memudahkan model dalam menggeneralisasi atau beradaptasi dengan berbagai data dan juga dapat membedakan antara *output* (Sharma, 2017). Terdapat beberapa *activation function* yang tersedia dalam *Non-linear* yaitu Sigmoid, Tanh, ReLU dan Leaky ReLU.

- **ReLU (*Rectified Linear Unit*)**

Aktivasi ReLU (*Rectified Linear Unit*) merupakan lapisan aktivasi pada model dengan mengaplikasikan fungsi $f(x) = \max(0, x)$.



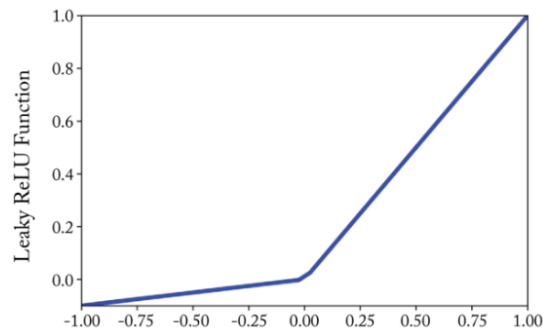
Gambar 2. 7 Fungsi Aktivasi ReLU (Sharma, 2017)

Fungsi aktivasi ini membuat seluruh nilai yang bernilai kurang dari nol (negatif) dijadikan 0 dan yang bernilai lebih dari 0 maka itu adalah nilai aktivasinya sendiri (Setiawan, 2022). Grafik untuk ReLU bisa dilihat di Gambar 2.7.

Menjadi masalah dengan ReLU adalah semua nilai yang hasilnya negatif akan menjadi nol yang akan menurunkan kemampuan model untuk menyesuaikan atau *training* dari data dengan benar. Dikarenakan neuron dengan nilai 0 yang berarti mati dengan jenis input apapun yang diberikan, sehingga performa *neural network* bisa terganggu (Sharma, 2017).

- **Leaky ReLU**

Untuk mengatasi masalah pada ReLU, maka dapat menggunakan fungsi Leaky ReLU. Fungsi ini memberikan mengubah nilai negatif menjadi nilai yang tidak sama dengan 0. Grafik dari Leaky ReLU bisa dilihat di Gambar 2.8 (Sharma, 2017).



Gambar 2. 8 Fungsi Aktivasi Leaky ReLU (Arif, 2022)

2.13 Algoritma You Only Look Once

Salah satu teknik yang digunakan dalam mendeteksi objek adalah metode *You Only Look Once* atau yang disingkat dengan YOLO dimana algoritma ini akan lebih cepat dalam melakukan identifikasi. Target dari metode YOLO adalah untuk dapat mendeteksi objek secara *real time*. Secara umum, YOLO melakukan kalkulasi *bounding box* dengan satu skala *feature map*.

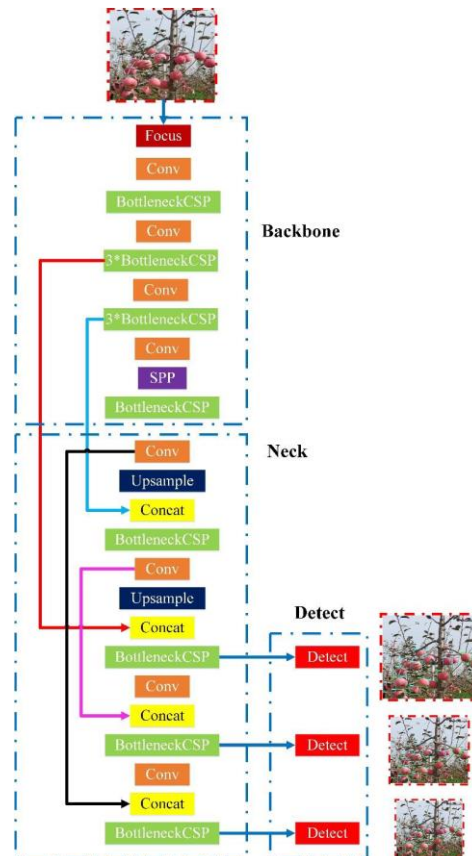
Tahapan YOLO dimulai dengan input gambar dengan ditandai sejumlah $S \times S$ grid box, kemudian setiap grid box akan membentuk sejumlah B “*bounding box*” yang tetap dengan berbagai ukuran yang bervariasi, jika pada *bounding box* juga terdapat objek, maka *grid box* yang membentuk *bounding box* tersebut bertanggung jawab untuk mendeteksi objek tersebut. Ketika pendeteksian pada *grid box* selesai, maka akan menerapkan *Non Max Supression* yang digunakan untuk

menentukan saat terdapat banyak *bounding box* terdeteksi untuk satu objek yang sama (Batubara & Awangga, 2020).

YOLOv5 merupakan salah satu produk dari seri arsitektur YOLO. Dengan keakuratan tinggi dalam model dan kecepatan inferensi yang cepat dengan kecepatan deteksi tercepat mencapai 140 *frames per second*. Di sisi lain, ukuran file bobot model jaringan deteksi target YOLOv5 kecil, yang hampir 90% lebih kecil dari YOLOv4 menunjukkan bahwa model YOLOv5 ini sangat cocok untuk diterapkan pada *embedded device* dalam deteksi *real-time*. Oleh karena itu, keunggulan jaringan YOLOv5 adalah akurasi deteksi yang tinggi, karakteristik yang ringan dan kecepatan deteksi yang cepat secara bersamaan.

Arsitektur YOLOv5 memiliki empat arsitektur yaitu: YOLOv5s, YOLOv5m, YOLOv5l dan YOLOv5x. Perbedaan utama di antara arsitektur itu adalah jumlah modul ekstraksi fitur dan kernel konvolusi di lokasi spesifik jaringan. Ukuran model dan jumlah parameter model dalam empat arsitektur itu meningkat secara bergantian (Yan dkk., 2021).

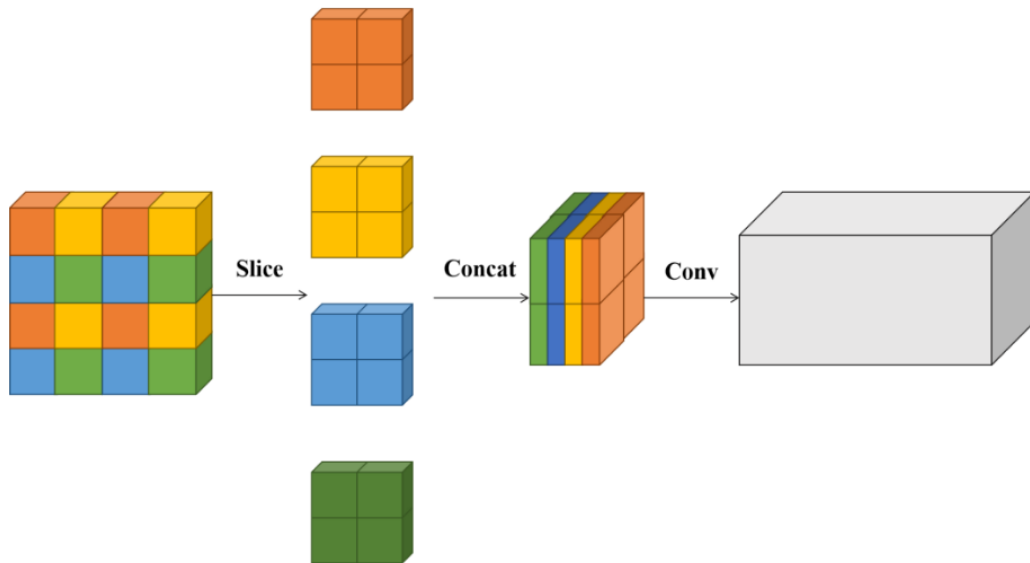
YOLOv5 menggunakan strategi pengoptimalan algoritma yang sangat baik di bidang *Convolutional Neural Network*, seperti pada *bounding box anchors*, *mosaic data augmentation*, *cross stage partial network* dan sebagainya. Di dalam arsitektur YOLOv5 terdiri dari empat bagian: *input*, *backbone*, *neck* dan *output*. Seperti yang digambarkan pada Gambar 2.9 di bawah ini



Gambar 2. 9 Arsitektur YOLOv5 (Nguyen dkk., 2021)

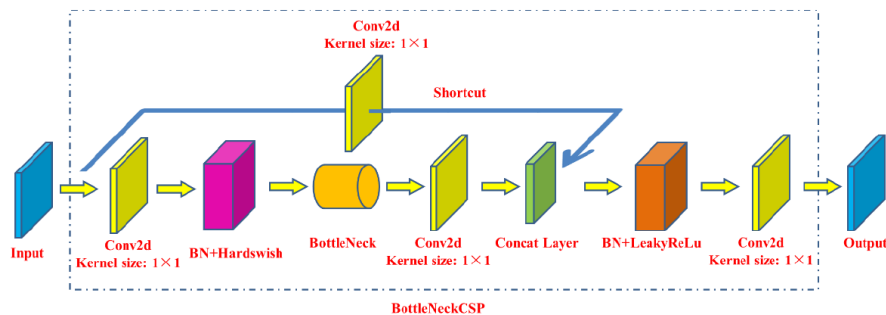
Untuk bagian *input* berisi *preprocessing data* termasuk augmentasi data *mosaic* dan pengisian citra adaptif. Dalam beradaptasi dengan kumpulan data yang berbeda, YOLOv5 memasukkan perhitungan *anchor frame* pada input, sehingga dengan otomatis dapat mengatur ukuran *anchor frame* saat datasetnya berubah (Li dkk., 2022).

Backbone network bekerja dalam *convolutional neural network* yang menggabungkan berbagai gambar dan membentuk fitur dalam gambar. Pada layer pertama dari *backbone* adalah modul *Focus*, dimana modul ini dirancang untuk mengurangi parameter dan ruang memori GPU serta mempercepat kecepatan *training*. Pertama, gambar input 3 channel masuk dan disegmentasikan menjadi empat irisan (*slice*). Selanjutnya masuk ke operasi *concat* digunakan untuk menggabungkan empat bagian dan kemudian masuk ke lapisan konvolusi untuk memperluas fitur (Yan dkk., 2021). Seperti pada Gambar 2.10 di bawah ini:

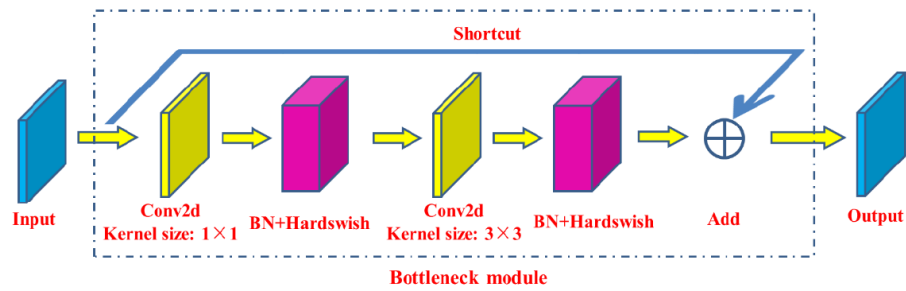


Gambar 2. 10 Struktur Focus pada YOLOv5 (Jing dkk., 2022)

Pada bagian *Conv* merujuk pada operasi konvolusi untuk pemrosesan fitur yang terdiri dari *convolutional layer*, *batch normalization*, dan *activation layer* menggunakan *Leaky ReLU*. Setelah dari bagian *Conv*, masuk ke lapisan ketiga yaitu *BottleneckCSP* (pada Gambar 2.11) yang dirancang untuk mengekstrak fitur-fitur dalam gambar dengan lebih baik. Input awal dari *BottleneckCSP* masuk ke dalam dua cabang operasi konvolusi dan melalui modul *Bottleneck* dan lapisan *Conv2d*. Dalam modul *Bottleneck* pada Gambar 2.12, input dikirim ke dua cabang ke *Conv2d*, *batch normalization* dan fungsi *activation*, kemudian output akhir melakukan penambahan output dan input awal. Kembali ke *BottleneckCSP*, keluaran dari dua cabang digabungkan dengan operasi *concat*. Dan di output modul ini hasil *concat* melewati BN, Leaky ReLU dan *Conv2d* (Yan dkk., 2021).

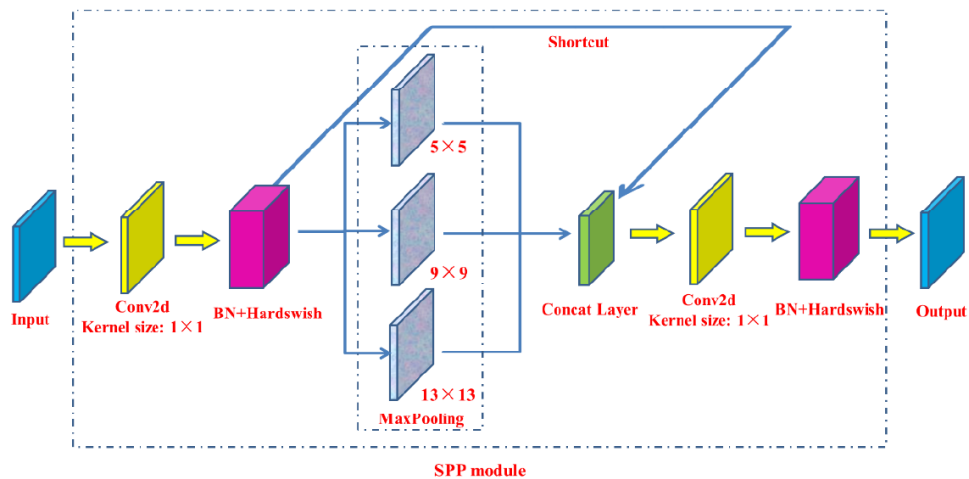


Gambar 2. 11 Struktur Module BottleneckCSP



Gambar 2. 12 Struktur Module Bottleneck

Lapisan kesembilan dari *Backbone network* adalah modul SPP (*Spatial Pyramid Pooling*) (pada Gambar 2.13) yang digunakan untuk meningkatkan bidang penerimaan jaringan dengan mengubah berbagai ukuran *feature map* menjadi *feature vector* dengan ukuran tetap. Dalam SPP dimulai dengan input dikirimkan ke lapisan konvolusi kemudian masuk ke tiga lapisan *Maxpooling* secara paralel dihubungkan secara mendalam. Pada akhir modul ini, *feature map* dapat diperoleh setelah melewati lapisan konvolusi (Yan dkk., 2021).



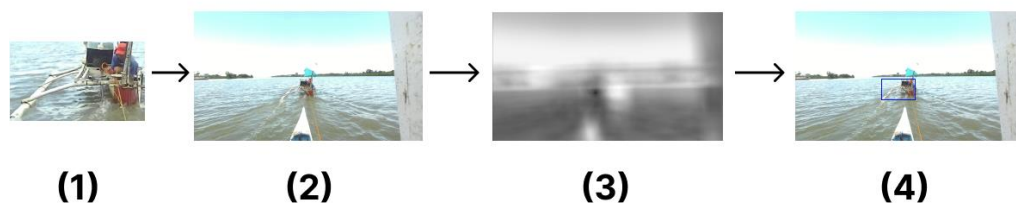
Gambar 2. 13 Struktur Module SPP

Bagian *neck network* memiliki serangkaian lapisan agregasi fitur dari penggabungan fitur gambar, yang digunakan untuk menghasilkan FPN (*Feature Pyramid Structures*) dan kemudian outputnya ditransmisikan ke *detect network*. *Neck network* ini dapat membantu model dalam mengidentifikasi target dalam skala yang berbeda.

Akhir dari arsitektur YOLOv5 ini adalah bagian *detect network* digunakan untuk menerapkan *anchor boxes* pada *feature map* dari lapisan sebelumnya dan mengeluarkan vektor dengan probabilitas kategori objek target, skor objek, dan posisi kotak pembatas yang mengelilingi objek. Pada bagian ini, terdapat tiga lapisan deteksi untuk mendeteksi objek gambar dengan ukuran yang berbeda dan kemudian diberi *bounding box* dan label ke objek yang diprediksi pada gambar (Yan dkk., 2021).

2.14 Template Matching

Template Matching adalah sebuah teknik dalam pengolahan citra digital untuk menemukan bagian-bagian terkecil dalam pencocokan gambar dengan template gambar. *Template Matching* bekerja dengan menggeser gambar template di atas gambar input dan membandingkan gambar *template* dan gambar input.



Gambar 2. 14 Proses dari Template Matching

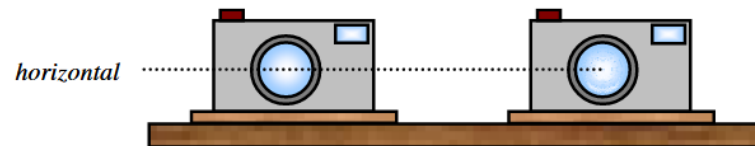
Berdasarkan pada Gambar 2.14, pada bagian pertama mengambil potongan gambar sebagai *template*, dan masuk di bagian kedua dimana *template* melakukan *scanning* pada gambar yang ingin dicari. Hasil *scanning* akan mengembalikan gambar *grayscale* pada bagian 3, titik cerah yang ada di gambar tersebut menandakan bahwa di titik itu adalah posisi yang tepat untuk gambar *template*. Kemudian hasil akhir dari *template matching* adalah nilai *bounding box* dalam piksel.

2.15 Stereoscopic Measurement

Stereoscopic merupakan teknik yang digunakan dalam merekam dan mewakili gambar berbentuk 3D. Dengan teknik ini dapat menciptakan ilusi

kedalaman menggunakan dua gambar yang diambil pada posisi yang sedikit berbeda. Teknik pengukuran *stereoscopic* (*Stereoscopic Measurement*) dapat digunakan dalam mencari jarak dengan memperoleh dua gambar yang berisi target objek (Dawood dkk., 2017).

Teknik *stereoscopic* dapat dilakukan dengan sepasang kamera, mirip dengan mata kita sendiri. Hal yang paling penting dalam mengambil sepasang gambar *stereoscopic* adalah kamera harus disejajarkan secara horizontal dan gambar harus diambil pada saat yang sama seperti pada Gambar 2.15 (Mrovlje & Vran, 2008).



Gambar 2. 15 Ilustrasi teknik *stereoscopic*

2.16 Penelitian Terkait

Salah satu penelitian yang telah menerapkan metode *You Only Look Once* (YOLO) untuk mendeteksi kapal adalah penelitian dari Sung-Jun Lee, dkk. dengan menggunakan data gambar kapal dari Pascal VOC dan Singapore Maritime Dataset dimana hasil yang didapatkan dari *Intersection Over Union* sebesar 66.69% dan Recall yang dihasilkan sebesar 77.12%.

Kemudian dalam penelitian dari Sarma Emani, dkk. melakukan pendeteksian jarak menggunakan *stereo vision* dengan metode *Single Shot Detector* sebagai pendeteksi objek dan ditempatkan diatas mobil buggy. Dalam penelitian ini mengambil jarak aktual 1 sampai 3.9 meter, dan hasil prediksi jarak yang didapatkan tidak lebih dari 10 cm.

Untuk mendeteksi jarak lebih jauh, penelitian dari Yasir Dawood Salman, dkk. menggunakan metode *Stereo Image Measurement* dalam mengukur jarak dengan *stereo vision* yaitu menggunakan dua kamera sebagai input dan hasil dalam penelitian ini adalah dapat mendeteksi objek dari 8,5 sampai 160 meter.

Tabel 2. 1 Penelitian Terkait

Judul	Penulis	Penerbit/ Tahun	Metode	Hasil
Simple Convolutional Neural Network on Image Classification	Tianmei Guo, dkk.	Institute of Electrical and Electronics Engineers (IEEE) /2017	Convolutional Neural Network	Menggunakan Dataset MNIST yang berisi gambar grayscale dari tulisan tangan angka 0-9, tingkat error yang didapat adalah 0.66%
Distance Measurement For Self-Driving Cars Using Stereo Camera	Yasir Dawood Salman, dkk.	International Conference of Computing & Informatics /2017	Stereo Image Measurement	Dengan menggunakan metode stereo image measurement, dalam eksperimen berhasil mendeteksi sampai 160m
Image-based Ship Detection and Classification for Unmanned Surface Vehicle Using Real-Time	Sung-Jun Lee, dkk.	International Society of Offshore and Polar Engineers (ISOPE)/2018	You Only Look Once (YOLO)	Hasil yang didapatkan dari verifikasi untuk Recall sebesar 77.12% dan IOU sebesar 66.69%

Object Detection Neural Networks				
Obstacle Detection and Distance Estimation for Autonomous Electric Vehicle Using Stereo Vision and DNN	Sarma Emami, dkk.	Springer Nature Singapore Pte Ltd./2019	Single Shot Detector (SSD), Deep Neural Network (DNN)	Hasil error estimasi jarak didapatkan tidak lebih dari 10cm saat percobaan dari jarak 100cm, 140cm, 190cm, 240cm, 290cm dan 390cm