

SKRIPSI

**SISTEM DETEKSI DAN ESTIMASI JARAK TANGGA
SECARA REALTIME BERBASIS ANDROID UNTUK
PENYANDANG TUNANETRA**

Disusun dan diajukan oleh:

HERLINA ANWAR

D121171009



DEPARTEMEN TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS HASANUDDIN

MAKASSAR

2022

LEMBAR PENGESAHAN SKRIPSI
SISTEM DETEKSI DAN ESTIMASI JARAK TANGGA SECARA
REALTIME BERBASIS ANDROID UNTUK PENYANDANG TUNANETRA

Disusun dan diajukan oleh

HERLINA ANWAR
D121171009

Telah dipertahankan di hadapan Panitia Ujian yang dibentuk dalam rangka Penyelesaian Studi Program Sarjana Program Studi Teknik Informatika Fakultas Teknik Universitas Hasanuddin pada tanggal 21 Juli 2022 dan dinyatakan telah memenuhi syarat kelulusan.

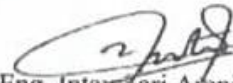
Menyetujui,

Pembimbing Utama,

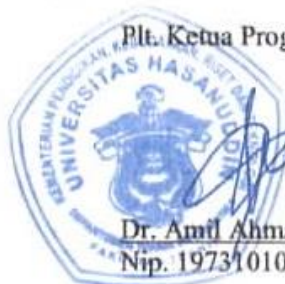
Pembimbing Pendamping,



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



Dr. Eng. Intan Bari Areni, S.T., M.T.
Nip. 19750203 200012 2 002



Plt. Ketua Program Studi,

Dr. Amil Ahmad Ilham, S.T., M.IT.
Nip. 19731010 199802 1 002

PERNYATAAN KEASLIAN

Yang bertanda tangan di bawah ini :

Nama : Herlina Anwar
NIM : D121171009
Program Studi : Teknik Informatika
Jenjang : S1

Menyatakan dengan ini bahwa Skripsi dengan judul

SISTEM DETEKSI DAN ESTIMASI JARAK TANGGA SECARA REALTIME BERBASIS ANDROID UNTUK PENYANDANG TUNANETRA

adalah karya saya sendiri dan tidak melanggar hak cipta pihak lain. Apabila di kemudian hari Skripsi karya saya ini terbukti bahwa sebagian atau keseluruhannya adalah hasil karya orang lain yang saya pergunakan dengan cara melanggar hak cipta pihak lain, maka saya bersedia menerima sanksi.

Makassar, 21 Juli 2022

Yang menyatakan,



HERLINA ANWAR

KATA PENGANTAR

Puji dan syukur penulis panjatkan atas kehadiran Allah SWT karena berkat Rahmat dan Karunia-Nya sehingga Tugas Akhir yang berjudul “SISTEM DETEKSI DAN ESTIMASI JARAK TANGGA SECARA REALTIME BERBASIS ANDROID UNTUK PENYANDANG TUNANETRA” ini dapat diselesaikan sebagai salah satu syarat dalam menyelesaikan jenjang Strata-1 pada Departemen Teknik Informatika Fakultas Teknik Universitas Hasanuddin.

Penulis menyadari bahwa dalam penyusunan dan penulisan skripsi ini tidak lepas dari bantuan, bimbingan serta dukungan dari berbagai pihak, dari masa perkuliahan sampai dengan masa penyusunan tugas akhir, sangatlah sulit untuk menyelesaikan tugas akhir ini. Oleh karena itu, penulis dengan senang hati menyampaikan terima kasih kepada:

1. Tuhan Yang Maha Esa atas semua berkat, karunia serta pertolongan-Nya yang telah diberikan kepada kami di setiap langkah dalam pembuatan program hingga penulisan laporan skripsi ini.
2. Orang tua penulis, Bapak Muh. Anwar dan Ibu Rahmawati yang tidak pernah lelah dalam mendidik dan memberikan dukungan, doa, dan semangat kepada penulis.
3. Bapak Dr. Indrabayu, S.T., M.T., M.Bus.Sys., selaku pembimbing 1 dan Ibu Dr. Eng. Intan Sari Areni, S.T., M.T., selaku pembimbing II yang selalu menyediakan waktu, tenaga, pikiran dan perhatian yang luar biasa untuk mengarahkan penulis dalam penyusunan tugas akhir.

4. Bapak Dr. Amil Ahmad Ilham, S.T., M.IT., selaku ketua Departemen Teknik Informatika atas segala bimbingan dan dukungan selama masa perkuliahan.
5. Ibu Elly Warni, S.T., M.T., selaku Dosen Pembimbing Akademik penulis yang selalu membimbing dan menyediakan waktu, tenaga, dan perhatiannya selama masa perkuliahan.
6. Segenap Dosen dan Staf Departemen Teknik Informatika Fakultas Teknik Universitas Hasanuddin yang telah banyak membantu semasa perkuliahan dan dalam menyelesaikan tugas akhir.
7. Para teman-teman dan kakak-kakak AIMP *research group* Unhas yang telah memberikan begitu banyak bantuan selama penelitian, pengambilan data dan diskusi *progress* penyusunan tugas akhir.
8. Teman-teman RECOGN17ER atas dukungan, bantuan, semangat yang diberikan selama ini.
9. Kepada kak Yuliani, Taslinda, Nublan, Jum, dan Ojan yang telah memberikan dukungan dan bantuan kepada penulis untuk menyelesaikan tugas akhir.
10. Para sahabat serta seluruh pihak yang tidak sempat disebutkan dan tanpa sadar telah menjadi inspirasi serta banyak membantu penulis dalam menyelesaikan tugas akhir ini.

Akhir kata, penulis menyadari bahwa tugas akhir ini masih jauh dari kata sempurna, oleh karenanya diharapkan segala bentuk saran serta masukan

yang membangun dari berbagai pihak. Semoga tugas akhir ini dapat memberikan sumbangsih dan manfaat besar bagi kepentingan bersama.

Gowa, 21 Juli 2022

Penulis

ABSTRAK

Tunanetra merupakan kondisi seseorang yang mengalami gangguan atau hambatan dalam indra penglihatannya. Tercatat oleh Kementerian Ketenagakerjaan dan Transmigrasi RI tahun 2017 menyatakan jumlah penyandang tunanetra di Indonesia sebanyak 2.137.923. Berdasarkan Jumlah tersebut, penyandang tunanetra merupakan jumlah penyandang disabilitas terbanyak di Indonesia. Pengenalan kondisi jalan sangat penting bagi penyandang tunanetra. Umumnya penyandang tunanetra menggunakan tongkat untuk membantu dalam mengenali jalan. Namun, tongkat kurang akurat dalam mendeteksi objek rintangan dan navigasi. Salah satu objek yang dapat membahayakan keselamatan penyandang tunanetra adalah tangga. Oleh karena itu, peneliti merancang sebuah aplikasi untuk mendeteksi dan mengestimasi jarak tangga secara *realtime* berbasis android yang akan mengeluarkan suara peringatan adanya tangga naik pada jarak 2, 3, 4, dan 5 meter atau tangga turun pada jarak 2 meter. Adapun metode yang digunakan dalam mendeteksi yakni *Single Shot Multibox Detector (SSD) MobileNet*, sedangkan metode untuk mengestimasi jarak tangga yakni *Pinhole Camera Model*. Akurasi yang didapatkan dalam mendeteksi tangga dengan melakukan pengujian terhadap tangga naik dan tangga turun di dua belas lokasi berbeda yakni akurasi tangga naik pada jarak 2, 3, 4 dan 5 meter sebesar 97,78%, 99,33%, 99,13%, dan 97,58%, sedangkan akurasi tangga turun pada jarak 2 meter sebesar 80,49%. Adapun untuk estimasi jarak pada tangga naik menghasilkan nilai *error* sebesar 0,21, 0,14, 0,13, 0,11 pada masing-masing jarak 2, 3, 4, dan 5 meter, sedangkan pada tangga turun menghasilkan nilai *error* sebesar 0,24 pada jarak 2 meter.

Kata Kunci: tunanetra, deteksi tangga, *realtime*, *Single Shot Multibox Detector (SSD) MobileNet*, *Pinhole Camera Model*

DAFTAR ISI

KATA PENGANTAR	ii
ABSTRAK	v
DAFTAR ISI.....	vi
DAFTAR GAMBAR	viii
DAFTAR TABEL.....	ix
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	4
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	4
1.5 Batasan Masalah	5
1.6 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA	7
2.1 Penyandang Tunanetra.....	7
2.2 Deteksi Objek	9
2.3 Pengolahan Citra.....	10
2.4 Computer Vision.....	11
2.5 Python	12
2.6 Single Shot Multibox Detector (SSD) MobileNet.....	13
2.7 Pinhole Camera Model	15
2.8 State of The Art.....	17
BAB III METODOLOGI PENELITIAN	20
3.1 Tahapan Penelitian.....	20
3.2 Waktu dan Lokasi Penelitian	21
3.3 Instrumen Penelitian	22

3.4	Teknik Pengambilan Data.....	22
3.5	Perancangan Implementasi Sistem	25
3.6	Analisis Kerja Sistem.....	37
BAB IV HASIL DAN PEMBAHASAN		39
4.1	Deteksi Tangga	39
4.2	Estimasi Jarak Tangga	48
BAB V KESIMPULAN & SARAN		55
5.1	Kesimpulan	55
5.2	Saran	56
DAFTAR PUSTAKA		57
LAMPIRAN.....		60

DAFTAR GAMBAR

Gambar 2.1 Arsitektur SSD	14
Gambar 2.2 Arsitektur VGG-16.....	15
Gambar 2.3 Proyeksi Objek pada Lubang Jarum Pinhole Camera Model	16
Gambar 2.4 Pinhole Camera Model.....	16
Gambar 3.1 Tahapan Penelitian	20
Gambar 3.2 Povie (Point of View)	23
Gambar 3.3 Pengukuran Jarak	23
Gambar 3.4 Tangga Turun Jarak 3 meter	24
Gambar 3.5 Alur Penelitian	26
Gambar 3.6 Gambar Masukan	27
Gambar 3.7 Resize Gambar	27
Gambar 3.8 Labeling Image.....	28
Gambar 3.9 Contoh Hasil Checkpoint	32
Gambar 3.10 Hasil Training.....	33
Gambar 3.11 Tahapan Pinhole Camera Model.....	34
Gambar 3.12 Perbandingan Jarak Aktual dan Jarak Acuan.....	35
Gambar 4.1 Frame Tangga Naik yang Terdeteksi	42
Gambar 4.2 Frame Tangga Turun yang Terdeteksi	42
Gambar 4.3 Frame Tangga Naik yang Tidak Terdeteksi.....	44
Gambar 4.4 Frame Tangga Turun yang Tidak Terdeteksi.....	44
Gambar 4.5 Hasil Estimasi Jarak Tangga Naik	48
Gambar 4.6 Hasil Estimasi Jarak Tangga Turun	48
Gambar 4.7 Frame Hasil Estimasi Jarak Tangga Naik	53
Gambar 4.8 Frame Hasil Estimasi Jarak Tangga Turun	54

DAFTAR TABEL

Tabel 2.1 State of The Art.....	18
Tabel 3.1 Contoh Isi File .CSV Data Latih.....	28
Tabel 3.2 Parameter Latih.....	29
Tabel 3.3 Konvolusi Layer Pada MobileNet	30
Tabel 3.4 Konvolusi Layer untuk Classification dan Localization.....	31
Tabel 3.5 Perbandingan Jarak Aktual dan Jarak Acuan.....	35
Tabel 4.1 Hasil Deteksi Tangga Naik	40
Tabel 4.2 Hasil Deteksi Tangga Turun	41
Tabel 4.3 Akurasi Deteksi Tangga Naik.....	46
Tabel 4.4 Akurasi Deteksi Tangga Turun.....	47
Tabel 4.5 Perhitungan Error Estimasi Jarak Tangga Naik.....	50
Tabel 4.6 Perhitungan Error Estimasi Jarak Tangga Turun.....	51

BAB I

PENDAHULUAN

1.1 Latar Belakang

Tunanetra merupakan kondisi seseorang yang mengalami gangguan atau hambatan dalam indra penglihatannya. Berdasarkan tingkat gangguannya tunanetra dibagi dua yaitu buta total (*total blind*) dan masih mempunyai sisa penglihatan (*low vision*) (Pertuni, 2022).

Menurut data Organisasi Kesehatan Dunia (WHO) pada tahun 2010 jumlah penyandang tunanetra di dunia mencapai 285 juta orang dan sebesar 39 juta dari angka tersebut telah mengalami kebutaan dan sisanya dalam kondisi penglihatan lemah. Diperkirakan angka tersebut akan meningkat secara signifikan kedepannya (WHO, 2010). Tercatat oleh Kementerian Ketenagakerjaan dan Transmigrasi RI menyatakan jumlah penyandang tunanetra di Indonesia sebanyak 2.137.923 jiwa. Berdasarkan jumlah tersebut, penyandang tunanetra merupakan jumlah penyandang disabilitas terbanyak di Indonesia (Dewi Poerwanti, 2017).

Pengenalan kondisi jalan sangat penting bagi penyandang tunanetra. Umumnya penyandang tunanetra menggunakan tongkat untuk membantu dalam mengenali jalan. Namun, tongkat kurang akurat dalam mendeteksi objek rintangan dan navigasi (Rahman, et al., 2018). Salah satu objek yang dapat membahayakan keselamatan penyandang tunanetra adalah tangga. Menurut Biro Sensus Amerika Serikat, di Amerika lebih dari 1 juta orang terluka karena aktivitas naik dan turun tangga. Total kerugian yang tercatat untuk luka non kematian akibat kecelakaan ketika naik dan turun tangga adalah \$ 92 miliar setiap tahunnya (Supriyadi, 2018).

Namun, seiring perkembangan teknologi, telah banyak penelitian yang dikembangkan dalam membantu tunanetra dalam mendeteksi objek rintangan secara akurat, salah satunya dengan pemanfaatan kamera dan pengolahan citra.

Al Kadafi dan Utaminingrum, menyajikan sistem untuk mendeteksi penghalang secara *realtime* berbasis *mobile* dengan menggunakan analisis blob. Sistem ini menggunakan metode *Connected Component Labeling* yang mampu mendeteksi halangan *indoor* baik pada waktu pagi hari, siang hari, sore hari, maupun malam hari dengan akurasi 81,25%. Objek yang dapat dideteksi yakni objek dalam ruangan (*Indoor*) seperti manusia, tempat sampah, dinding, pot bunga, dan pintu (Al Kadafi, 2017). Indrabayu dkk, juga menyajikan sistem untuk mendeteksi dan mengestimasi jarak posisi relatif kendaraan motor terparkir secara *realtime* dengan menggunakan metode *Single Shot Multibox Detector (SSD)* *MobileNet* dan *Pinhole Camera Model*. Sistem ini mendeteksi melalui kamera *smartphone* dengan sudut 70 derajat dan 75 derajat dengan akurasi tertinggi 83,26% pada jarak 2 meter.

Berdasarkan penelitian sebelumnya, pemanfaatan kamera *smartphone* telah banyak digunakan untuk mendeteksi dan mengestimasi jarak objek hambatan secara *realtime* untuk membantu tunanetra dalam mengenali jalan.

Berdasarkan penelitian, algoritma *Convolutional Neural Network (CNN)* memiliki kinerja terbaik dalam klasifikasi citra secara umum dan telah terbukti keakuratannya dalam klasifikasi gambar dan pendeteksian objek. CNN juga memperoleh tingkat kesalahan lebih rendah dalam kompetisi *ImageNet*. Berdasarkan pengembangan algoritma CNN, terdapat tiga algoritma yang memiliki

performansi tinggi, yaitu algoritma *RCNN Faster, You Only Look Once* (YOLO), dan *Single Shot Multibox Detector* (SSD) (Gianani et al., 2018).

Algoritma *Single Shot Multibox Detector* (SSD) *MobileNet* merupakan algoritma *Deep Learning* yang dianggap paling kompatibel dengan kapasitas memori dan pemrosesan data yang terbatas karena mampu menghasilkan model data dengan kecepatan, akurasi, dan kapasitas yang memadai meskipun menggunakan dataset dalam jumlah yang sedikit (Ryu and Kim, 2018). Metode SSD dengan menggunakan *MobileNet* dapat berjalan dengan komputasi ringan, sehingga dapat dijalankan secara *realtime* di perangkat seluler (Liu et al., 2016). Sedangkan model *Faster-RCNN* lebih berat secara komputasi, tetapi menghasilkan pendeteksian yang jauh lebih akurat (Huang et al., 2017).

Selain mendeteksi objek tangga menggunakan algoritma SSD, dilakukan pula estimasi jarak penyandang tunanetra terhadap tangga. Hal tersebut bertujuan untuk memudahkan penyandang tunanetra dalam mengestimasi langkah. Metode yang digunakan untuk mengestimasi jarak objek tangga adalah *Pinhole Camera Model*. Metode ini menyatakan rasio ukuran objek dalam piksel dan ukuran objek dalam dunia nyata setara dengan rasio antara panjang fokus dari lensa kamera dan jarak objek ke kamera (Parihar et al., 2017). Pada umumnya, metode ini digunakan pada aplikasi sistem *realtime* dengan pengambilan data menggunakan kamera (Mane and Yangandul, 2016). Hal ini sejalan dengan sistem yang diusulkan menggunakan kamera dengan pengaplikasiannya menggunakan *smartphone android*.

1.2 Rumusan Masalah

Berdasarkan latar belakang, maka rumusan masalah dalam tugas akhir ini adalah:

1. Bagaimana unjuk kerja pendeteksian objek tangga menggunakan SSD *MobileNet*?
2. Bagaimana unjuk kerja *Pinhole Camera Model* dalam mengestimasi jarak tangga?

1.3 Tujuan Penelitian

Tujuan dari tugas akhir ini adalah untuk:

1. Mengetahui unjuk kerja pendeteksian objek tangga menggunakan metode *SSD MobileNet*.
2. Mengetahui unjuk kerja *Pinhole Camera Model* dalam mengestimasi jarak tangga.

1.4 Manfaat Penelitian

Manfaat dari tugas akhir ini adalah:

1. Bagi penyandang tunanetra, sistem dapat digunakan untuk mendeteksi dan mengestimasi jarak tangga sehingga dapat membantu penyandang tunanetra ketika berjalan.
2. Bagi Institusi pendidikan, dapat digunakan sebagai referensi dalam pengembangan penelitian topik terkait penelitian pendeteksian dan estimasi jarak tangga.

1.5 Batasan Masalah

Yang menjadi batasan masalah dalam tugas akhir ini adalah:

1. Objek yang dideteksi yakni tangga dengan warna yang berbeda dan ukuran yang berbeda.
2. Pengambilan data menggunakan kamera *smartphone* yang dipasang pada *povie*.
3. Pengambilan data dilakukan di lingkungan Fakultas Teknik Universitas Hasanuddin, Mall, Masjid, Jembatan penyeberangan, Rumah Kayu, dan Toko.
4. Data diambil pada siang hari dengan jarak tangga naik yaitu 2, 3, 4, dan 5 meter sedangkan jarak tangga turun yaitu 2 meter.
5. Data diambil menggunakan kamera *smartphone* dengan tingkat kemiringan kamera yaitu 70 derajat.
6. Data latih berupa gambar dan data uji berupa video.

1.6 Sistematika Penulisan

Laporan Tugas Akhir ini dibagi dalam lima bab yang tersusun secara sistematis, yaitu:

BAB I PENDAHULUAN

Bab ini menjelaskan tentang latar belakang pengambilan topik penelitian sistem deteksi dan estimasi jarak tangga secara *realtime* berbasis android untuk penyandang tunanetra, disertai rumusan masalah, tujuan penelitian, batasan masalah, metode penelitian, dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Bab ini akan membahas mengenai penjelasan-penjelasan umum tentang teori-teori yang berkaitan dengan konsep dasar sistem deteksi dan estimasi jarak tangga secara *realtime* berbasis android.

BAB III METODOLOGI PENELITIAN

Bab ini membahas perancangan dan penerapan algoritma dalam sistem deteksi dan estimasi jarak tangga.

BAB IV HASIL DAN PEMBAHASAN

Bab ini berisi tentang hasil pengujian sistem deteksi dan estimasi jarak tangga serta pembahasan yang disertai tabel hasil penelitian.

BAB V PENUTUP

Bab ini berisi tentang kesimpulan yang didapatkan berdasarkan hasil penelitian yang telah dilakukan serta saran-saran untuk pengembangan lebih lanjut.

BAB II

TINJAUAN PUSTAKA

2.1 Penyandang Tunanetra

Tunanetra adalah individu yang memiliki hambatan dalam penglihatan. tunanetra dapat diklasifikasikan kedalam dua golongan yaitu: buta total (*blind*) dan *low vision*. Persatuan Tunanetra Indonesia (Pertuni) mendefinisikan tunanetra adalah mereka yang tidak memiliki penglihatan sama sekali (buta total) serta mereka yang masih memiliki sisa penglihatan tetapi tidak mampu menggunakan penglihatannya untuk membaca tulisan biasa berukuran 12 point dalam keadaan cahaya normal dari jarak normal meskipun dibantu dengan kaca mata (kurang awas / *low vision*) (Pertuni, 2022).

Seseorang dikatakan buta apabila mempergunakan kemampuan perabaan dan pendengaran sebagai indra utama dalam belajar. Mereka mungkin mempunyai sedikit persepsi cahaya atau persepsi bentuk atau sama sekali tidak dapat melihat (buta total). Orang seperti ini biasanya mempergunakan huruf *Braille* sebagai media membaca dan memerlukan latihan orientasi dan mobilitas (Rahardja, 2009)

Seseorang dikatakan menyandang *low vision* apabila masih memungkinkan memfungsikan indra penglihatannya dalam melakukan kegiatan sehari-hari. Indra utama yang dipergunakan dalam belajar adalah penglihatan dengan mempergunakan alat bantu, baik yang direkomendasikan oleh dokter maupun bukan. Jenis huruf yang dipergunakan sangat bervariasi tergantung pada sisa penglihatan dan alat bantu yang dipergunakannya. Latihan orientasi dan mobilitas

diperlukan oleh penyandang *low vision* untuk mempergunakan sisa penglihatannya (Rahardja, 2009).

Penyebab ketunanetraan sangat bervariasi tergantung lokasi geografis, status sosial, ekonomi, dan usia. Secara umum sebetulnya bisa dicegah dan diatasi. Trachoma merupakan penyebab utama timbulnya kebutaan di negara-negara berkembang. Banyak organisasi yang berhubungan dengan kesehatan mempunyai program pencegahan kebutaan. Mereka bekerja di perkampungan dan daerah-daerah miskin dengan tujuan untuk memberikan penyuluhan kepada masyarakat tentang kebersihan, kesehatan, dan akses untuk memperoleh pengobatan. Diabetes, glaukoma, dan katarak merupakan penyebab umum kebutaan di negara-negara barat. Hal ini terjadi karena usia harapan hidup mereka lebih panjang dari generasi sebelumnya; usia berhubungan dengan penurunan daya penglihatan (Rahardja, 2009).

Perilaku sosial secara tipikal dikembangkan melalui observasi terhadap kebiasaan dan kejadian sosial serta menirunya. Perbaikan biasanya dilakukan melalui penggunaan yang berulang-ulang dan bila diperlukan meminta masukan dari orang lain yang berkompeten. Karena tunanetra mempunyai keterbatasan dalam belajar melalui pengamatan dan menirukan, penyandang tunanetra sering mempunyai kesulitan dalam melakukan perilaku sosial yang benar. Sebagai akibat dari ketunanetraannya yang berpengaruh terhadap keterampilan sosial, penyandang tunanetra harus mendapatkan pembelajaran yang langsung dan sistematis dalam bidang pengembangan persahabatan, menjaga kontak mata atau orientasi wajah, penampilan postur tubuh yang baik, mempergunakan gerakan tubuh dan ekspresi

wajah dengan benar, mengekspresikan perasaan, menyampaikan pesan yang tepat pada waktu melakukan komunikasi, serta mempergunakan alat bantu yang tepat (Rahardja, 2009).

2.2 Deteksi Objek

Deteksi objek adalah teknik visi komputer untuk menemukan objek dalam gambar atau video. Algoritma deteksi objek biasanya memanfaatkan pembelajaran mesin atau pembelajaran mendalam untuk menghasilkan hasil yang bermakna. Ketika 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 (Aningtiyas, 2020).

Cara kerja deteksi objek yakni menempatkan keberadaan objek dalam gambar dan menggambar kotak pembatas di sekitar objek itu. Ini biasanya melibatkan dua proses, yaitu mengklasifikasikan jenis objek, dan kemudian menggambar kotak di sekitar objek itu. Klasifikasi gambar dan skenario deteksi objek terlihat serupa. Secara umum, klasifikasi adalah mengklasifikasikan gambar ke dalam kategori tertentu. Sedangkan objek deteksi adalah mengidentifikasi lokasi objek dalam gambar, dan misalnya menghitung jumlah instance suatu objek (Aningtiyas, 2020).

2.3 Pengolahan Citra Digital

Pengolahan citra digital merupakan metode yang digunakan untuk mengolah citra pada komputer sehingga dapat menghasilkan gambar sesuai dengan yang dibutuhkan. Misalnya terdapat sebuah citra digital berwarna dengan ukuran 1280×720 piksel. Dengan pengolahan citra digital, gambar tersebut dapat diubah ukurannya menjadi 640×360 tanpa mengurangi kualitas gambar.

Citra digital merupakan sebuah larik (*array*) yang berisi nilai-nilai real maupun kompleks yang direpresentasikan dengan deretan bit tertentu. Suatu citra dapat didefinisikan sebagai fungsi $f(x,y)$ berukuran M baris dan N kolom, dengan x dan y adalah koordinat spasial, dan amplitudo f di titik koordinat (x,y) dinamakan intensitas atau tingkat keabuan dari citra pada titik tersebut. Nilai pada suatu irisan antara baris dan kolom (pada posisi x, y) disebut *dengan picture element, image element, pels*, atau *pixel*. Istilah terakhir (*pixel*) adalah istilah yang paling sering digunakan pada citra digital (Sutoyo, 2009).

Citra bergerak (*moving images*) adalah rangkaian citra diam yang ditampilkan secara beruntun sehingga memberi kesan pada mata kita sebagai gambar yang bergerak. Setiap citra di dalam rangkaian itu disebut *frame*. Gambar-gambar yang tampak pada film layar lebar atau televisi pada hakikatnya terdiri dari ratusan hingga ribuan *frame*. Meskipun sebuah citra kaya informasi, namun seringkali citra yang kita miliki mengalami penurunan mutu (*degradasi*), misalnya mengandung cacat atau derau (*noise*), warnanya terlalu kontras, kurang tajam, kabur (*blurring*), dan sebagainya. Tentu saja citra semacam ini menjadi lebih sulit diinterpretasi karena informasi yang disampaikan oleh citra tersebut menjadi berkurang. Agar

citra yang mengalami gangguan mudah diinterpretasikan dengan baik oleh manusia maupun mesin, maka citra tersebut perlu dimanipulasi menjadi citra lain yang kualitasnya lebih baik (Syawaluddin, 2016).

2.4 Computer Vision

Computer vision adalah suatu proses transformasi atau perubahan dari data yang berasal dari kamera video maupun foto/gambar kedalam sebuah hasil keputusan ataupun sebuah presentasi yang baru, dimana hasil dari kegiatan transformasi tersebut memiliki kepentingan untuk mencapai suatu tujuan. Data yang dimasukkan kedalam kegiatan transformasi tersebut memungkinkan untuk memiliki beberapa informasi yang terkontekstual seperti halnya sebuah foto/gambar yang didalamnya terdapat berbagai objek. Dengan demikian akan didapatkan keputusan-keputusan yang akan diambil pada gambar (Aningtiyas, 2020).

Secara garis besar, *computer vision* adalah sebuah teknologi mesin yang mampu mengenali objek yang diamati. Kemampuan untuk mengenali ini merupakan kombinasi dari pengolahan citra dan pengenalan pola (Basuki, 2016). Pengolahan citra adalah proses awal dalam *computer vision* untuk menghasilkan citra yang lebih baik atau lebih mudah diinterpretasikan, sedangkan pengenalan pola adalah proses 12 identifikasi objek pada citra. Proses-proses dalam *computer vision* secara garis besar dapat dibagi menjadi (Basuki, 2016):

1. Proses mengakuisisi citra digital (*Image Acquisition*)
2. Proses pengolahan citra (*Image Processing*)

3. Proses analisis data citra (*Image Analysis*)

4. Proses pemahaman data citra (*Image Understanding*)

Computer vision (visi komputer) dapat didefinisikan dengan pengertian pengolahan citra yang dikaitkan dengan akuisisi citra, pemrosesan, klasifikasi, dan pencakupan keseluruhan, pengambilan keputusan yang diikuti pengidentifikasian citra. Inti dari teknologi *Computer vision* adalah untuk menduplikasi kemampuan penglihatan manusia ke dalam benda elektronik sehingga benda elektronik dapat memahami dan mengerti arti dari gambar yang dimasukkan (Prabowo, 2018).

2.5 Python

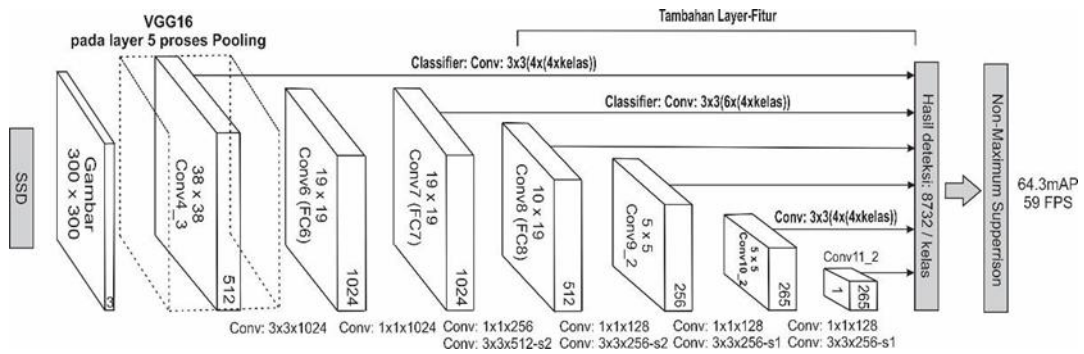
Python adalah suatu bahasa pemrograman tingkat tinggi yang berfokus pada keterbacaan kode-kodenya. Python merupakan salah satu bahasa pemrograman yang sangat mudah dipahami. Kode-kode python memiliki kemiripan dengan bahasa manusia. Sebagai contohnya perintah `print()` yang digunakan untuk mencetak suatu string ke console. Bahasa tersebut pastinya sangat akrab di telinga kita (Aulia, 2021)

Secara sederhana, sebuah komputer hanya dapat mengeksekusi program yang ditulis dalam bentuk bahasa mesin. Oleh karena itu, jika suatu program ditulis dalam bentuk bahasa tingkat tinggi, maka program tersebut harus diproses dulu sebelum bisa dijalankan dalam komputer. Hal ini merupakan salah satu kekurangan bahasa tingkat tinggi yang memerlukan waktu untuk memproses suatu program sebelum program tersebut dijalankan. Akan tetapi, bahasa tingkat tinggi mempunyai banyak sekali keuntungan. Bahasa tingkat tinggi mudah dipelajari,

mudah ditulis, mudah dibaca, dan tentu saja mudah dicari kesalahannya. Bahasa tingkat tinggi juga mudah diubah portabel untuk disesuaikan dengan mesin yang menjalankannya. Hal ini berbeda dengan bahasa mesin yang hanya dapat digunakan untuk mesin tersebut. Dengan berbagai kelebihan ini, maka banyak aplikasi ditulis menggunakan bahasa tingkat tinggi. Proses mengubah bentuk bahasa tingkat tinggi ke tingkat rendah dalam bahasa pemrograman ada dua tipe, yakni interpreter dan compiler (Utami, 2004).

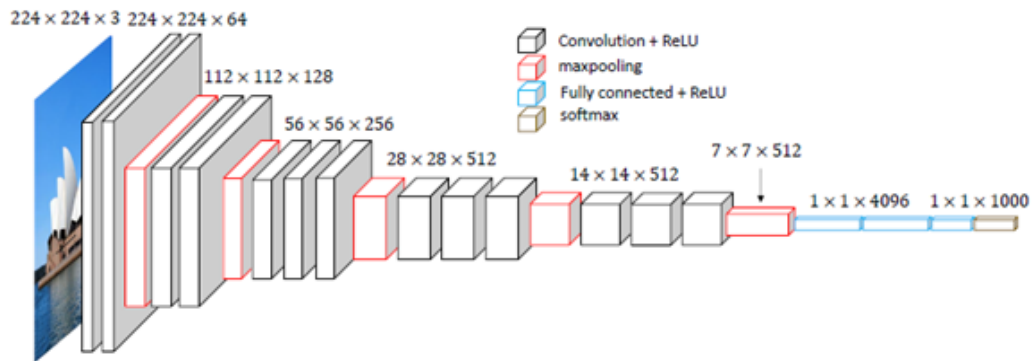
2.6 Single Shot Multibox Detector (SSD) MobileNet

Single Shot Multibox Detector (SSD) merupakan algoritma pengembangan *Deep Neural Network* (DNN) yang bertujuan untuk mendeteksi objek. SSD menjalankan proses *classification* objek dan proses *localization* secara bersamaan dalam satu proses *feed-forward*. Algoritma SSD hanya membutuhkan citra masukan dan *ground truth* dari objek selama proses latih (pelatihan). SSD menggunakan pendekatan berbasis *feed-forward convolutional network* yang nantinya menghasilkan kumpulan *bounding box* (*bbox*) dan skor untuk menentukan kategori dari objek pada masing-masing *bounding box* tersebut. Jaringan dasar yang digunakan adalah VGG16 dan diikuti oleh beberapa jaringan konvolusional. Arsitektur Algoritma SSD dapat dilihat pada Gambar 2.1 berikut:



Gambar 2.1 Arsitektur SSD

SSD tersusun dalam beberapa proses konvolusi. Bagian awal SSD menggunakan arsitektur dari VGG16 (Ranalli et al., 2018) yang merupakan salah satu algoritma klasifikasi citra berkualitas tinggi. VGG16 pada SSD terdiri dari 5 lapisan konvolusi yang masing-masing menghasilkan *feature map* yang menjadi masukan pada lapisan setelahnya. Pada layer konvolusi keempat, *feature map* yang dihasilkan digunakan sebagai masukan untuk proses *localization* dan *classification* pertama. Setiap *cell* (lokasi, koordinat) pada *feature map* digunakan untuk membentuk 4 objek prediksi. Setiap prediksi terdiri dari satu *bbox* dan $N + 1$ skor untuk tiap kelas (tambahan 1 untuk kelas “bukan objek”) dimana N adalah jumlah kelas. SSD menggunakan filter konvolusi berukuran 3×3 untuk melakukan prediksi di tiap *cell* pada *feature map*. Tiap filter tersebut menghasilkan *output* dengan *channel* $N + 1 + 4$ dimana tambahan 4 adalah lokasi dari *bbox* (x, y, w, h). Misalkan pada lapisan konvolusi keempat dari VGG16 menghasilkan *feature map* dengan ukuran $38 \times 38 \times 512$. Dengan menggunakan filter dengan ukuran $3 \times 3 \times 512$, SSD akan menghasilkan output sebesar $38 \times 38 \times 4 \times (N+1+4)$.



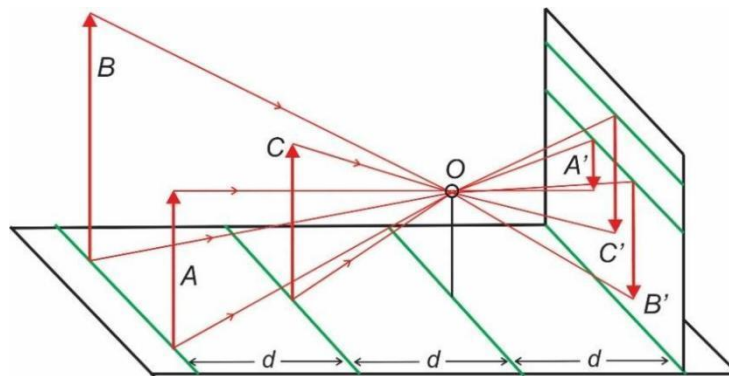
Gambar 2.2 Arsitektur VGG-16

Gambar 2.2 merupakan arsitektur dari lapisan VGG, dimana arsitektur SSD dibangun diatas arsitektur VGG-16, tetapi membuang lapisan yang terhubung sepenuhnya, Alasan VGG-16 digunakan sebagai jaringan dasar adalah karena kinerjanya yang kuat dalam tugas-tugas klasifikasi gambar berkualitas tinggi sehingga dapat meningkatkan hasil. Alih-alih VGG asli sepenuhnya terhubung lapisan, satu set lapisan konvolusional tambahan (dari conv6 dan seterusnya) ditambahkan, sehingga memungkinkan untuk mengekstrak fitur pada berbagai skala dan secara progress mengurangi ukuran *input* ke setiap lapisan berikutnya (Fermansah, 2019).

2.7 Pinhole Camera Model

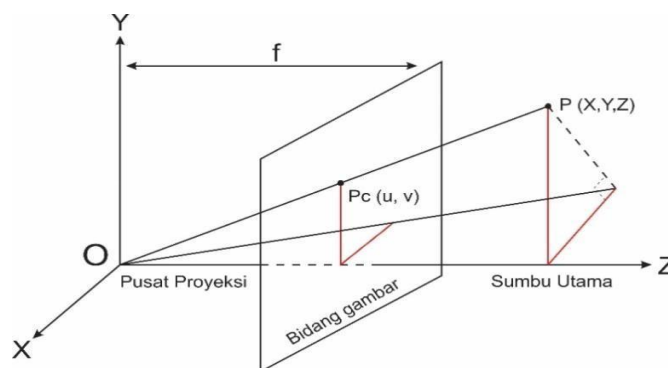
Pinhole Camera Model dimodelkan sebagai kotak tertutup dengan lubang kecil yang dilubangi dengan pin di salah satu sisinya. Sinar cahaya yang memasuki kotak hanya melalui lubang jarum ini membentuk gambar terbalik di sisi berlawanan dari kotak yang disebut bidang gambar. Gambar ini dibentuk oleh perpotongan sinar cahaya yang melewati lubang jarum dengan bidang

gambar, seperti yang diilustrasikan pada Gambar 2.3. Objek A , B , dan C pada jarak d yang melalui lubang jarum O diproyeksikan sebagai A' , B' , dan C' dengan perbandingan ukuran dan jarak yang sama dengan objek sebenarnya.



Gambar 2.3 Proyeksi Objek pada Lubang Jarum Pinhole Camera Model (Majumder and Gopi, 2018)

Pinhole Camera Model menggunakan prinsip *similar triangle*. Seperti yang terlihat pada Gambar 2.3, objek hasil proyeksi akan membentuk segitiga terbalik yang sama dengan yang dihasilkan pada objek sebenarnya. Berdasarkan prinsip tersebut, metode *Pinhole Camera Model* dapat digunakan untuk mengestimasi jarak pada gambar. Jarak objek sebenarnya dapat diproyeksikan pada gambar seperti yang diilustrasikan pada Gambar 2.4.



Gambar 2.4 Pinhole Camera Model (Majumder and Gopi, 2018)

Pada Gambar 2.4, titik O merupakan pusat proyeksi kamera. P merupakan objek tiga dimensi dengan X , Y , dan Z masing-masing adalah lebar, tinggi, dan jaraknya terhadap kamera. Hasil proyeksi P yang ditangkap oleh kamera adalah P_c sebagai citra dua dimensi u , v yang masing-masing merupakan koordinat x dan y objek dalam *pixel*. Citra hasil proyeksi berada pada jarak f dari titik O dengan f adalah *focal length* kamera. Karena adanya perpotongan antara bidang gambar dengan sumbu utama Z , maka digunakan prinsip *similar triangle*, sehingga diperoleh persamaan berikut:

$$\frac{f}{Z} = \frac{u}{X} = \frac{v}{Y}$$

2.8 State of The Art

Beberapa penelitian terkait deteksi objek banyak dilakukan sebelumnya dengan beberapa metode dan algoritma sebagai berikut:

Tabel 2.1 State of The Art

No	Judul	Peneliti	Penerbit/Tahun	Metode	Hasil
1	Sistem Deteksi Lubang pada Pedestrian dengan Teknik Pengolahan Citra	Indrabayu, dkk	Jurnal Penelitian <i>Engineering</i> , 23 (2), pp.117-120 / 2019	Deteksi: <i>Threshold+Blob Analysis</i> dan Metode HVS Estimasi: <i>Pinhole Camera Model</i>	Lubang terdeteksi dengan menggunakan mono kamera dengan tingkat akurasi rata-rata 88,91% menggunakan metode <i>Threshold+Blob Analysis</i> . Dan 86,82% untuk metode HVS.
2	Deteksi Tangga Turun, Tangga Naik, dan Lantai menggunakan <i>Gray Level Co-Occurrence Matrix</i> dan <i>K-Nearest Neighbors</i> berbasis <i>Raspberry Pi</i>	Ester Nadya Fiorentina Lumban Gaol, dkk	Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer Vol. 5, No. 2, Februari 2021, hlm. 675-680 / 2021	Metode: <i>Gray Level Co-occurrence matrix</i> (GLCM) dan <i>K-Nearest Neighbors</i> (K-NN)	Dengan penggunaan metode GLCM dengan 4 fitur (<i>contrast, dissimilarity, energy, dan homogeneity</i>), nilai jarak ketetangaan piksel ($d = 4$, $teta=90^\circ$ dan K-NN dengan $k=3$ pada sistem, diperoleh rata-rata akurasi sebesar 93,33% dan rata-rata waktu komputasi 1551ms. Dengan output suara.
3	<i>Obstacle Detection in Indoor Environment for Visually Impaired Using Mobile Camera</i>	Sami ur Rahman, dkk	<i>Journal of Physics: Conference Series</i> / 2017	Segmentasi: ROI Deteksi: MSE	Menyajikan sistem <i>realtime</i> dengan akurasi 96% menggunakan kamera <i>smartphone</i> dengan tingkat kemiringan 40 derajat untuk mendeteksi hambatan di lingkungan <i>indoor</i>
4	Deteksi Objek Penghalang Secara <i>Realtime</i> Berbasis <i>Mobile</i> Bagi Penyandang Tunanetra Menggunakan Analisis Blob.	Achmad Jafar Al Kadafi, dkk	Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer e-ISSN, 2548, p.964X / 2018	Segmentasi: <i>Thresholding</i> Deteksi: <i>Connected Component Labeling</i>	Sistem yang dibangun mampu untuk mendeteksi halangan dalam kondisi <i>indoor</i> baik pada waktu pagi hari, siang hari, sore hari, maupun malam hari dengan nilai akurasi sebesar 81,25%.

Lanjutan Tabel 2.1

No	Judul	Peneliti	Penerbit/Tahun	Metode	Hasil
5	<i>Detection and Distance Estimation against Motorcycles as Navigation Aids for Visually-impaired People</i>	Indrabayu, dkk	<i>2019 12th International Conference on Information & Communication Technology and System (ICTS) (pp. 224-228). IEEE / 2019</i>	Deteksi: <i>SSD Mobilenet</i> Estimasi: <i>Pinhole Camera Model</i>	Sepeda motor terdeteksi dan akurasi tertinggi dari perkiraan posisi relatif adalah 83,26% pada jarak 2 meter dengan rata-rata RMSE dari estimasi jarak adalah 0,11. Dengan output suara