

## Deteksi Objek Menggunakan Metode *Single Shot Multibox Detector* Pada Alat Bantu Tongkat Tunanetra Berbasis Kamera

Samratul Fuady\*, Nehru, Gina Anggraeni

Program Studi Teknik Elektro, Fakultas Sains dan Teknologi, Universitas Jambi

\*Corresponding e-mail: sfuady@unja.ac.id

**Abstrak** Penyandang tunanetra mengalami kesulitan untuk bernavigasi dikarenakan keterbatasan penginderaan yang dialami. Pada penelitian ini telah dirancang sebuah tongkat yang dapat membedakan antara halangan berupa manusia, hewan, dan benda berbasis kamera. Pemrosesan dilakukan dengan Raspberry Pi dengan input kamera webcam dan indikator berupa buzzer dan vibrator. Proses ekstraksi fitur dilakukan dengan *deep learning* menggunakan *library tensorflow* dan pengolahan citra dengan metode *Single Shot MultiBox Detector* (SSD). Pengujian dilakukan terhadap objek manusia, hewan (kucing), dan benda (kursi dan meja) untuk kondisi dalam dan luar ruangan dan didapatkan nilai akurasi sebesar 92% , sensitifitas 83%, dan spesifisitas 100% .

**Kata Kunci** : Tongkat tunanetra, Deteksi objek dan SSD

**Abstract** *Blind people have difficulty in navigating due to the limited sensing they are capable of. In this research, we design a stick tool that can distinguish objects in the form of humans, animals and inanimate object based on camera. Processing is carried out with the Raspberry Pi with a webcam camera as input and indicators in the form of a buzzer and vibrator. The feature extraction process is carried out by deep learning using the tensorflow library and image processing using the Single Shot MultiBox Detector (SSD) method. Tests were carried out on human objects, animals (cats), and inanimate objects (chairs and tables) for indoor and outdoor conditions and obtained an accuracy of 92%, a sensitivity of 83%, and a specificity of 100%.*

**Keywords** : *Blind stick, Object detection, and SSD*

### 1. PENDAHULUAN

Dalam kehidupan sehari-hari penyandang tunanetra mengalami kesulitan untuk beraktifitas dikarenakan keterbatasan penginderaan yang di alami. Untuk melakukan aktifitas dalam bernavigasi penyandang tunanetra menggunakan tongkat untuk melakukan orientasi mobilitas dan mendeteksi adanya halangan saat berjalan [1].

Seiring dengan semakin berkembangnya teknologi, beberapa penelitian telah dilakukan untuk meningkatkan kemampuan deteksi dan efektifitas penggunaan alat bantu tongkat tunanetra. Penelitian oleh [2][3] menggunakan sensor ultrasonik untuk mendeteksi adanya objek atau halangan dan memberikan indikator yang dapat dikenali oleh penggunaanya. Namun, pendeteksian halangan menggunakan sensor ultrasonik tidak dapat membedakan jenis objek yang ada di depannya[4], sehingga pengguna tidak memiliki informasi yang cukup untuk memberikan respon yang sesuai.

Karena itu, pada penelitian ini dilakukan perancangan alat bantu tongkat tunanetra berbasis kamera. Secara garis besar prinsip kerja sistem ini adalah dapat mendeteksi objek yang dianggap sebagai penghalang menggunakan kamera yang dipasang pada tongkat, kemudian kamera akan menangkap citra digital video dan mengirimkannya ke Raspberry Pi untuk dilakukan pemrosesan selanjutnya. Penerapan *Computer vision* dengan metode *Single Shot Multibox Detector* (SSD) digunakan pada penelitian ini untuk mengidentifikasi objek berupa halangan secara *real-time*.

### 2. TINJAUAN PUSTAKA

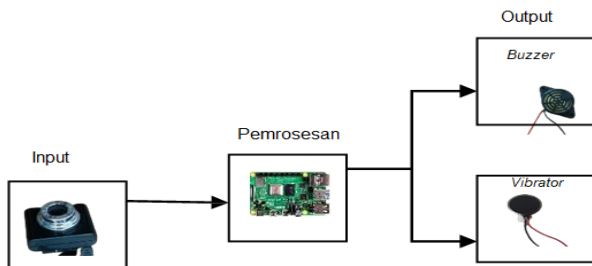
Pengolahan citra digital merupakan pemrosesan gambar maupun gambar bergerak berdimensi dua melalui komputer digital. Secara umum hal ini dibagi menjadi dua proses yaitu memperbaiki kualitas suatu gambar sehingga dapat lebih mudah diinterpretasi oleh manusia dan mengolah informasi yang terdapat pada suatu gambar untuk pengenalan objek secara otomatis[5].

*Single Shot Detector* (SSD) adalah sebuah metode untuk mengenali atau mendeteksi sebuah object pada suatu gambar dengan menggunakan *single deep neural network* dan salah satu algoritma deteksi object yang paling populer karena kemudahan implementasi, serta akurasi yang baik relatif terhadap komputasi yang dibutuhkan [6]. Metode *Single Shot Detector* (SSD) ini termasuk kedalam deteksi object secara *real time*. Arsitektur SSD termasuk kedalam jenis *Convolutional Neural Network* (CNN) [7], yang merupakan salah satu jenis Neural Network yang biasa digunakan pada data image. Arsitektur dari CNN dibagi menjadi 2 bagian besar, *Feature Extraction Layer* dan *Convolutional Layer*. Dimana pada bagian *Feature Extraction Layer* ini adalah melakukan *encoding* dari sebuah image menjadi *features* yang merepresentasikan gambar tersebut. Sedangkan bagian *Convolutional Layer* terdiri dari neuron yang tersusun sedemikian rupa sehingga membentuk sebuah filter dengan panjang dan tinggi (*pixels*).

Dataset pada penelitian menggunakan model pendeteksian objek SSDLite-MobileNet-v2 terkuantisasi yang dilatih dari set data Common Object in Context

Detection Challenge (COCO) [8] dan dikonversi yang telah disimpan pada ruang penyimpanan Raspberry Pi. Dan dataset tersebut digunakan untuk klasifikasi objek. Dataset ini telah meliputi 90 jenis objek, mulai dari manusia, hewan (kucing, anjing, kuda, dsb.), serta benda (mobil, kursi, meja, dsb.)

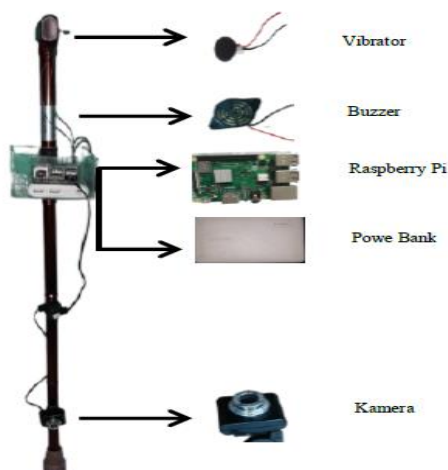
### 3. METODOLOGI



Gambar 1. Gambaran umum sistem

Sistem tongkat pendeteksi halangan bagi tunanetra terdiri dari tiga bagian, yaitu input, pemrosesan, dan output. Input pada sistem ini yaitu berasal dari kamera yang digunakan mengambil citra digital. Kamera akan menangkap citra digital secara real time dan mengirimkannya ke Raspberry Pi. Pada bagian pemrosesan dilakukan pengolahan citra yang diproses menggunakan metode *Single Shot Detector* (SSD) pada Raspberry Pi, kemudian menghasilkan gambar yang sudah memiliki deteksi objek. Setelah melakukan pemrosesan terhadap citra, sistem akan memberikan output berupa suara dari *buzzer* dan getar dari *vibrator* ketika berhasil mendeteksi objek yang ada di depan pengguna.

Tongkat yang digunakan terbuat dari aluminium dengan tinggi tongkat 95 cm berat 600 gr dilengkapi dengan kamera 5.0 MP dengan sudut kamera *wide angle*, *power bank* berkapasitas 10400 mAH, raspberry PI, buzzer dan vibrator dengan rancangan seperti gambar 2.



Gambar 2. Desain tongkat tunanetra berbasis kamera

Untuk perangkat lunak dalam mengimplementasikan sistem ini digunakan OpenCV untuk mengolah video secara *real-time*. Setiap frame dari video tersebut akan diolah oleh Raspberry Pi dengan *Single Shot MultiBox Detector* (SSD). Python IDLE digunakan sebagai *environment* pemrograman yang dilakukan. Dengan metodel *Single Shot MultiBox Detector* (SSD) citra asli dari kamera akan diekstraksi oleh layer konvolusi, kemudian Neural Network akan mengklasifikasikan objek sesuai dengan dataset yang sudah tersedia. Setelah mendapatkan hasil dari pengolahan citra, *buzzer* dan *vibrator* akan mengeluarkan peringatan atau indikator.

Klasifikasi objek yang dilatih pada dataset Common Object in Context Detection Challenge dikelompokkan menjadi tiga kelas berbeda, yaitu

1. Manusia
2. Hewan
3. Benda

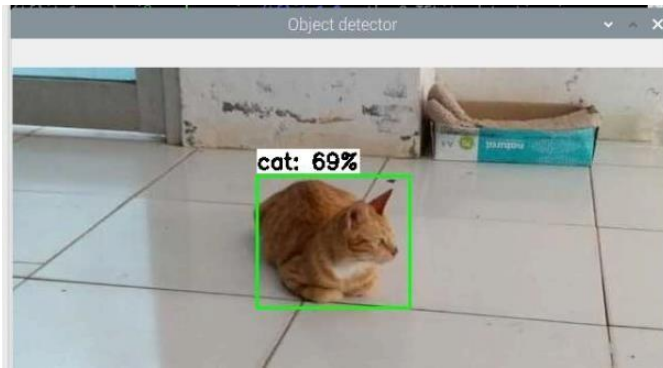
Tahapan proses pengolahan citra menggunakan metode *Single Shot MultiBox Detector* (SSD) dibagi menjadi beberapa tahap. Proses awalnya yaitu konvolusi untuk mendapatkan feature pada sebuah citra. Proses ini membentuk sebuah filter dengan panjang dan tinggi (*pixels*). Citra akan *resize* ke 300x300 *pixels* dengan kedalaman 3 *channels* yaitu RGB (*Red, Green, Blue*). Lapisan konvolusi ini dibentuk dengan menjalankan filter yang digunakan untuk menentukan pola apa yang akan dideteksi. Selanjutnya dilakukan konvolusi kembali sampai mendapatkan bagian terkecil dari sebuah citra.

Selanjutnya adalah tahap *pooling* yang menggunakan metode *max pooling* yang bertujuan untuk menentukan nilai maksimum pada setiap nilai input yang telah dikonvolusi dengan filter.

Terakhir, adalah proses *fully connected layer* yaitu tahap yang menghubungkan konvolusi dan *pooling*. Pada proses ini piksel yang dianggap sebagai halangan akan menjadi sebuah output yang terdiri dari salah satu kelas label. Hasil klasifikasi tersebut akan menentukan bagian mana saja pada setiap piksel yang memiliki pola manusia, hewan dan benda.

Pada proses ini, pengambilan citra dilakukan secara terus menerus (*real time*) untuk melakukan pendeteksian objek. Citra yang telah di ambil dari kamera selanjutnya akan diolah melalui pengolahan citra dan hasilnya berupa objek deteksi, dalam pengolahan citra ini menggunakan *Tensorflow*.

Terdapat 3 proses untuk mendapatkan deteksi objek. Tahap pertama, menentukan kotak pembatas untuk mengukur seberapa jauh kotak terkait jaringan yang diprediksi dan koordinat objek yang terdeteksi. Tahap kedua yaitu mendapatkan *class* yang dapat diklasifikasikan objek yang terdeteksi. Dan tahap terakhir adalah mendapatkan score nilai probabilitas dari objek yang terdeteksi.



Gambar 3. Tampilan dari pengolahan citra



Gambar 4. Pengujian deteksi manusia



Gambar 5. Pengujian deteksi benda

## 4. HASIL DAN PEMBAHASAN

### 4.1. Fungsionalitas Pendeteksian Objek

Pada tahap ini akan dilakukan uji coba pendeteksian citra terhadap objek berupa manusia dan benda. Pengujian dilakukan sebanyak tiga kali untuk masing-masing jarak, dan dilakukan di dalam dan diluar ruangan. Hasil pengujian dapat dilihat pada tabel 1 dan 2.

Tabel 1. Hasil pengujian deteksi manusia

Jarak	Tingkat Keyakinan	
	Dalam Ruangan	Luar Ruangan
30cm	59%	-
50cm	69%	-
75cm	-	65%
1 m	-	69%
2 m	74%	68%
3 m	74%	73%
4 m	78%	74%
5 m	73%	73%
10 m	-	73%
15 m	-	68%

- : tidak terdeteksi

Tabel 2. Hasil pengujian deteksi benda

Jarak	Tingkat Keyakinan	
	Dalam Ruangan	Luar Ruangan
30cm	-	-
50cm	58%	62%
75cm	64%	68%
1 m	61%	61%
2 m	71%	53%
3 m	74%	-
4 m	67%	-
5 m	65%	-
10 m	-	-
15 m	-	-

- : tidak terdeteksi

Dari dua pengujian yang dilakukan di dalam ruangan dan di luar ruangan terlihat bahwa jarak deteksinya di luar ruangan lebih jauh di bandingkan di dalam ruang. Tingkat ketelitian dari setiap jarak berbeda-beda dikarenakan pada setiap jarak posisi objek dan intensitas cahaya berbeda-beda. Untuk mendapatkan hasil maksimal kualitas dari kamera harus diperhatikan karena semakin baik resolusi kamera, citra yang di hasilkan semakin baik.

Dari hasil pengujian ini didapatkan jarak kerja efektif untuk sistem ini untuk deteksi manusia adalah 2m-5m dan untuk deteksi benda adalah 50cm-2m.

### 4.2. Pengujian Prototipe

Pengujian ini dilakukan untuk mengetahui keseluruhan sistem dapat berkerja sesuai dengan yang diharapkan dan mengetahui kemampuan sistem dalam mendeteksi adanya halangan, Pengujian ini dilakukan dengan metode *confusion matrix* [9].

Pengujian dilakukan di dalam ruangan dan diluar ruangan. Jarak yang digunakan adalah dalam rentang jarak efektif yang didapat pada uji fungsionalitas sebelumnya. Untuk masing-masing tempat dilakukan 30 kali pengujian, masing-masing 5 kali untuk kondisi terdapat objek target dan 5 untuk tidak terdapat objek target. Objek yang menjadi target adalah manusia, benda (kursi), dan hewan (kucing). Pengamatan dilakukan dengan mencatat respon yang diberikan sistem dan dikategorikan dalam 4 nilai, yaitu true positive (TP), true negative (TN), false positive (FP), dan false negative (FN). Hasil yang Hasil yang diperoleh dapat dilihat pada tabel 3 dan 4.





Gambar 5. Pengujian prototipe di dalam ruangan



Gambar 6. Pengujian prototipe di luar ruangan

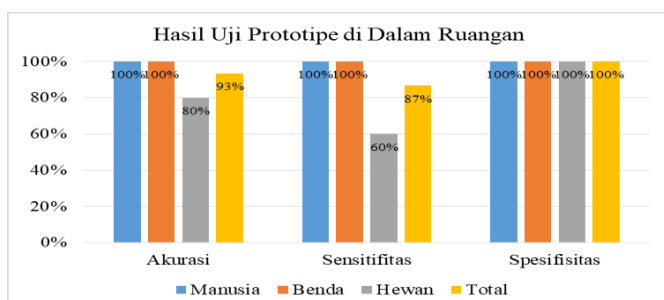
Tabel 3. Hasil pengujian prototipe di dalam ruangan

	Manusia	Benda	Hewan	Total
TP	5	5	3	13
TN	5	5	5	15
FP	0	0	0	0
FN	0	0	2	2
Total	10	10	10	30

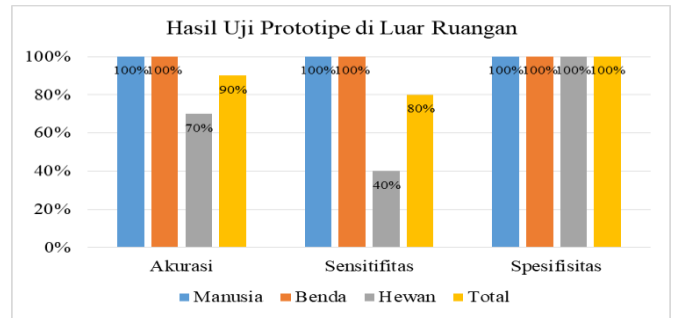
Tabel 4. Hasil pengujian prototipe di luar ruangan

	Manusia	Benda	Hewan	Total
TP	5	5	2	12
TN	5	5	5	15
FP	0	0	0	0
FN	0	0	3	3
Total	10	10	10	30

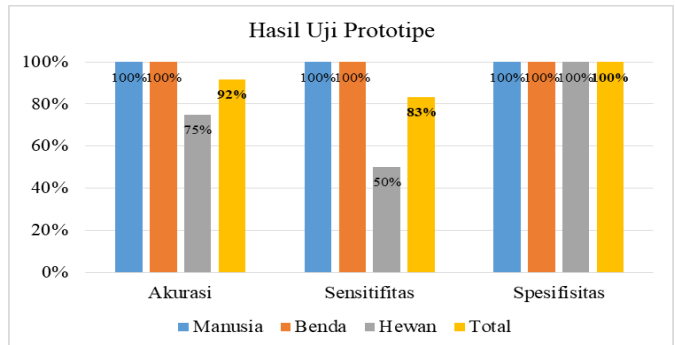
Dari pengujian ini, terlihat bahwa pendeteksian manusia dan benda dapat dilakukan dengan sangat baik. Namun, untuk pendeteksian hewan, beberapa kali mengalami kesalahan, baik di dalam maupun di luar ruangan. Untuk melihat kinerja dari sistem, dari data ini dihitung akurasi, sensitifitas, dan spesifisitas. Hasilnya dapat dilihat pada gambar 7-9. Secara keseluruhan, akurasi sistem berada di angka 92%, kesalahan banyak terjadi saat pendeteksian hewan. Untuk sensitifitas adalah sebesar 83% dan spesifisitasnya 100%.



Gambar 7. Hasil pengujian prototipe di dalam ruangan



Gambar 8. Hasil pengujian prototipe di luar ruangan



Gambar 9. Hasil pengujian prototipe keseluruhan

## 5. KESIMPULAN

Pada penelitian ini telah dirancang alat bantu tongkat tunanetra untuk pendeteksian objek berbasis kamera. Metode yang digunakan adalah Single Shot Multibox Detector karena metode ini mampu memproses citra secara real-time dengan tingkat akurasi yang cukup baik. Dari pengujian yang dilakukan terhadap objek manusia, hewan (kucing), dan benda (kursi dan meja) untuk kondisi dalam dan luar ruangan dan didapatkan nilai akurasi sebesar 92%, sensitifitas 83%, dan spesifisitas 100%. Akurasi identifikasi jenis halangan sangat dipengaruhi oleh kondisi citra saat diuji. Akurasi terbaik adalah citra dalam keadaan terang dan posisi yang tepat, sedangkan terendah adalah dalam keadaan gelap dan posisi tidak tepat.

## DAFTAR PUSTAKA

- [1] Andreas, Wisnu Wendanto. "Tongkat Bantu Tunanetra Pendeteksi Halangan Menggunakan Sensor Ultrasonik Berbasis Mikrokontroler Arduino." *Jurnal Ilmiah Go Infotech* 22, no. 1 (2016).
- [2] Y. Liu, Z. Gao, Z. Shao and G. Liu, "Intelligent ultrasonic detection of walking sticks for the blind," *2017 9th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, Targoviste (2017)
- [3] Bunnan, Shraddha, Gagan Pratap Singh, and S. P. Tondare. "Ultrasonic Blind Walking Stick for the Visually Impaired." *IJRET* 5 (2016)
- [4] A. Sen, K. Sen and J. Das, "Ultrasonic Blind Stick for Completely Blind People to Avoid Any Kind of Obstacles," *IEEE SENSORS*, (2018)

- [5] Thohari, A. N. A., dan Adhitama, R. Real-Time Object Detection For Wayang Punakawan Identification Using Deep Learning. Jurnal Infotel (2019)
- [6] Liu, Wei, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C. Berg. "Ssd: Single shot multibox detector." In *European conference on computer vision*, pp. 21-37. Springer, Cham, (2016)
- [7] Ning, Chengcheng, Huajun Zhou, Yan Song, and Jinhui Tang. "Inception single shot multibox detector for object detection." In *2017 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, pp. 549-554. IEEE. (2017)
- [8] Lin, Tsung-Yi, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Lawrence Zitnick. "Microsoft coco: Common objects in context." In *European conference on computer vision*, pp. 740-755. Springer, Cham, (2014)
- [9] Visa, Sofia, Brian Ramsay, Anca L. Ralescu, and Esther Van Der Knaap. "Confusion Matrix-based Feature Selection." *MAICS* 710 (2011)