

SISTEM KONTROL PERANGKAT BERBASIS INFRAMERAH MENGGUNAKAN SPEECH RECOGNITION DENGAN SPECTROGRAM DAN CONVOLUTIONAL NEURAL NETWORK BERBASIS MIKROKONTROLER

Irfan Muzakky Nurrizqy¹, Barlian Henryranu Prasetyo^{*2}, Rekyan Regasari Mardi Putri³

^{1,2,3}Universitas Brawijaya, Malang

Email: ¹muzakkynur@gmail.com, ²barlian@ub.ac.id, ³rekyan.rmp@ub.ac.id

^{*}Penulis Korespondensi

(Naskah masuk: 07 Januari 2023, diterima untuk diterbitkan: 26 September 2023)

Abstrak

Menurut data dari Biro Pusat Statistik (BPS), terdapat sebanyak 22,5 juta dari penduduk Indonesia merupakan penyandang disabilitas. Angka ini berjumlah sekitar lima persen dari keseluruhan penduduk Indonesia. Di zaman sekarang, kemajuan teknologi di seluruh dunia berkembang dengan pesat, sehingga muncul banyak hal yang dapat membantu menyederhanakan kehidupan semua orang, terutama penyandang disabilitas. Salah satu hal yang membantu penyandang disabilitas adalah munculnya perangkat pintar yang dapat dikendalikan menggunakan indra selain tangan, seperti suara. Penelitian ini bertujuan untuk mengembangkan sistem yang dapat mengendalikan perangkat inframerah dengan menggunakan suara sebagai input. Sistem tersebut akan dikembangkan menggunakan mikrokontroler dan metode *speech recognition* yang terdiri dari *spectrogram* dan CNN. Penelitian ini direncanakan untuk tujuan untuk membantu penyandang disabilitas dalam mengendalikan perangkat-perangkat di sekitar rumah. Hasil pengujian menunjukkan bahwa akurasi model CNN sebesar 93% dan akurasi percobaan terhadap pengguna sebesar 74,25%. Sistem ini juga dapat menjalankan proses *speech recognition* dengan waktu rata-rata 0,105 detik. Jarak optimal yang diperlukan antara pengguna dengan mikrofon adalah 30 cm dan jarak optimal yang diperlukan antara transmitter inframerah dengan perangkat yang dikendalikan adalah 30 cm.

Kata kunci: *Spectrogram, Speech Recognition, Convolutional Neural Network, Inframerah, Suara*

INFRARED BASED DEVICE CONTROL SYSTEM USING SPEECH RECOGNITION WITH SPECTROGRAM AND CONVOLUTIONAL NEURAL NETWORK BASED ON MICROCONTROLLER

Abstract

According to data from the Central Bureau of Statistics (BPS), around 22.5 million of Indonesia's population are people with disabilities. This number amounts to about five percent of Indonesia's total population. In the present day, where technology advances are rapidly developing all around the world, there have been many things that can help simplify the lives of everyone in the world, especially people with disabilities. One thing that helps people with disabilities is the emergence of smart devices that do not need to be controlled using hands but can use other senses such as sound. This research aims to develop a system that can control infrared devices using sound as input. The system will be developed using microcontrollers and speech recognition methods consisting of spectrogram and CNN. This research is conducted with the goal of helping people with disabilities in controlling devices around the house. Testing results show that the accuracy of the CNN model is 93% and the accuracy of trials on users is 74.25%. The system can also run the speech recognition process with an average time of 0.105 seconds. The optimal distance required between the user and microphone is 30 cm and the optimal distance required between the infrared transmitter and the controlled device is 30 cm.

Keywords: *Spectrogram, Speech Recognition, Convolutional Neural Network, Infrared, Speech*

1. PENDAHULUAN

Menurut data dari Biro Pusat Statistik (BPS), pada tahun 2020 terdapat sebanyak 22,5 juta dari penduduk Indonesia merupakan penyandang

disabilitas. Angka ini berjumlah sekitar lima persen dari keseluruhan penduduk Indonesia. Di zaman sekarang, kemajuan teknologi di seluruh dunia berkembang dengan pesat, sehingga muncul banyak

hal yang dapat membantu menyederhanakan kehidupan semua orang, terutama penyandang disabilitas. Salah satu hal yang membantu penyandang disabilitas adalah munculnya perangkat pintar yang dapat dikendalikan menggunakan selain organ motorik.

Salah satu perangkat yang sering digunakan untuk memudahkan kehidupan penyandang disabilitas adalah perangkat rumah pintar. Pengaplikasian rumah pintar dapat dilakukan dengan beberapa cara, contohnya rumah pintar berbasis aplikasi android (Ramlee et al., 2013), rumah pintar berbasis *smart assistant* (Mtshali & Khubisa, 2019), rumah pintar berbasis *bluetooth* (Ramlee et al., 2012), dsb. Penelitian sebelumnya menunjukkan bahwa perancangan rumah pintar berbasis suara akan memakan biaya yang cukup besar.

Beberapa peralatan Nirkabel seperti Wireless LAN, Inframerah, Bluetooth, merupakan contoh komunikasi secara nirkabel yang sering digunakan sebagai protokol komunikasi antar perangkat dalam teknologi smarthome. Komunikasi bluetooth lebih stabil dan memiliki jangkauan sinyal yang lebih panjang daripada inframerah. Fitur-fitur ini membuatnya ideal untuk teknologi yang ditujukan untuk kenyamanan seperti earbud, headset, dan pengontrol game. Di sisi lain, inframerah lebih baik untuk komunikasi jarak pendek dan skala kecil, seperti pengontrol peralatan rumah tangga.

Untuk tujuan itu, penelitian ini adalah merancang perangkat pintar yang bertugas mengendalikan perangkat elektronik rumah berbasis inframerah menggunakan *speech recognition*. Pengenalan ucapan dikenali menggunakan fitur *spectrogram* dan algoritma CNN. Sistem yang diusulkan sebagai usulan dasar sebuah perangkat pintar smarthome berbasis embedded system.

Speech recognition atau pengenalan suara adalah teknologi yang memungkinkan komputer atau perangkat elektronik untuk menerima dan mengidentifikasi input suara dari pengguna. Teknologi ini banyak digunakan untuk memudahkan pengguna dalam mengontrol perangkat elektronik, mengisi formulir secara otomatis, atau membantu proses transkrip suara menjadi teks. *Speech recognition* menggunakan algoritma dan teknik yang berbeda-beda, salah satunya adalah algoritma Convolutional Neural Network (CNN) (Abdul Qayyum et al., 2019; Palaz et al., 2015).

Convolutional Neural Network (CNN) merupakan salah satu jenis Artificial Neural Network (ANN) yang khusus dirancang untuk tugas pengenalan gambar dan video. CNN terinspirasi dari struktur sistem visual manusia, yang memproses dan mengenali pola dalam gambar melalui penggunaan filter dan lapisan konvolusi. CNN terdiri dari input layer, beberapa lapisan konvolusi, pooling layer, dan output layer yang terhubung dengan jaringan neuron.

CNN dapat dilatih dengan data latih yang banyak untuk mengenali pola dalam data tersebut

dan memprediksi output yang tepat. CNN banyak digunakan dalam pengenalan gambar dan video, seperti pengenalan objek, wajah, dan teks. CNN juga dapat digunakan untuk *speech recognition* dengan mengubah suara menjadi gambar atau *spectrogram* dan kemudian melatih model dengan data latih suara.

Spectrogram adalah representasi visual dari suara yang menunjukkan distribusi frekuensi dan waktu suara tersebut. *Spectrogram* dibuat dengan menggunakan Fourier Transform yang mengubah sinyal suara dari domain waktu ke domain frekuensi, kemudian ditampilkan dalam bentuk grafik amplitudo terhadap frekuensi dan waktu. *Spectrogram* dapat digunakan dalam *speech recognition* (Badshah et al., 2017; Prasomphan, 2015) dengan memperlakukan suara sebagai gambar yang dapat dikenali oleh komputer melalui CNN yang dilatih untuk mengenali pola suara dan memprediksi output yang tepat. Ini membantu komputer atau perangkat elektronik untuk menerima dan mengidentifikasi input suara dari pengguna.

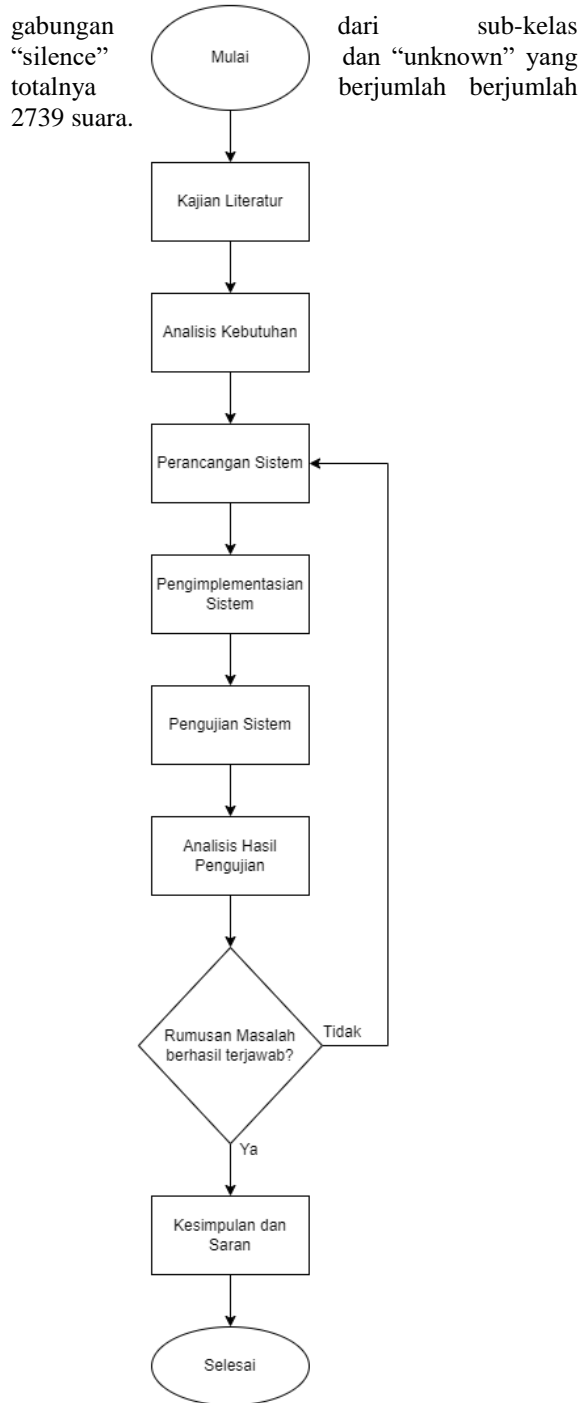
Ada beberapa hal yang perlu diperhatikan saat melakukan penelitian ini. Antara lain, akurasi *speech recognition* saat melakukan testing terhadap pengguna, akurasi model CNN yang buat dan pengaruh step terhadap akurasi model CNN, lama waktu komputasi yang diperlukan untuk melakukan *speech recognition*, jarak optimal yang diperlukan antara mikrofon dan pengguna, dan terakhir, jarak optimal antar *transmitter* inframerah dengan perangkat yang akan dikendalikan (Ajaegbu et al, 2020; Zou et al 2019).

2. METODE PENELITIAN

Tipe penelitian adalah penelitian implementatif pengembangan Penelitian implementatif berarti pengembangan sistem akan dilakukan penerapan di dunia nyata dari perangkat lunak dan juga perangkat keras. Penelitian pengembangan berarti sistem yang dibuat didasarkan dari penelitian sebelumnya untuk mencoba mengembangkan kekurangan dan kelebihan dari penelitian sebelumnya. Tahapan metode penelitian dapat dilihat pada **Error! Reference source not found..**

2.1 Teknik Pengumpulan dan Analisis Data

Data yang akan digunakan akan terbagi menjadi dua jenis data. Yang pertama yaitu data untuk model dan data untuk pengujian. Data model berisi 4 kelas utama dan 2 sub-kelas selain kelas utama yang disatukan menjadi 1 kelas. Jadi total ada 5 kelas yang akan diklasifikasikan. Kelas utama terdiri dari "Nyala" berjumlah 518 suara. Kelas "Mati" berjumlah 523 suara. Kelas "Naik" berjumlah 449 suara. Jumlah dataset dari kelas "Turun" berjumlah 417 suara. Kelas selain utama merupakan kelas "background" yang berisi



Gambar 1. Tahapan Metode Penelitian

Pada Gambar 2 dapat dilihat bahwa fase training model yang berisi langkah langkah yang dilakukan untuk membuat model CNN yang akan digunakan untuk mengklasifikasikan suara dari suara pengguna. Di bagian ini hasil tangkapan receiver inframerah dari remote perangkat inframerah yang akan dikendalikan juga akan dimasukkan ke dalam kode. Setelah model berhasil terbuat, akan lanjut ke bagian Proses Klasifikasi Suara. Bagian ini berisi tentang apa yang akan dilakukan untuk

mengklasifikasikan suara pengguna yang nantinya akan digunakan untuk mengendalikan perangkat inframerah. Suara akan diklasifikasikan kedalam kelas “Nyala”, “Mati”, “Naik”, “Turun”, dan “Background”.

3. PERANCANGAN DAN IMPLEMENTASI

3.1. Perancangan Perangkat Keras

Sistem dibuat berbasis mikrokontroler Seeed Studio Xiao Sense dengan mikrofon PDM sebagai *input*-nya. Sensor yang dipakai adalah *receiver* inframerah dan *transmitter* inframerah dan untuk daya memakai dua buah baterai 18650 yang dihubungkan ke step-down lm2596 untuk menurunkan voltase ke 5V sesuai dengan *input* daya mikrokontroler. Skematik perangkat keras dapat dilihat pada Gambar 3.

Tabel 1. Koneksi Pin antar komponen

Pin Baterai	Pin power converter
BAT+	IN+
BAT-	IN-
Pin power converter	Pin USB tipe C
OUT+	Power Cable +
OUT-	Power Cable -
Pin USB tipe C	Pin mikrokontroler
Konektor tipe C	Port input tipe C
Transmitter	Pin mikrokontroler
S	GPIO 1
+	VUSB (5V)
-	GND
Pin transmitter	Pin mikrokontroler
S	GPIO 0
Receiver	VUSB (5V)
+	VUSB (5V)
-	GND

3.2. Perancangan Perangkat Lunak

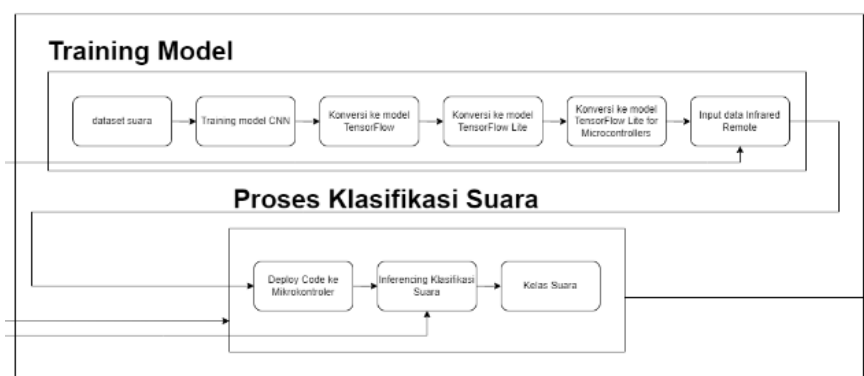
Perancangan perangkat lunak untuk memenuhi tujuan sistem dapat dilihat pada Gambar 4.

Pertama, sistem akan menginisialisasi variabel variabel yang akan digunakan. Lalu sistem akan mulai proses pengambilan suara yang. Setelah suara diambil, suara tersebut akan diberi fitur dan dijadikan *spectrogram*. Suara yang sudah diubah ke *spectrogram* akan diklasifikasikan dengan model CNN untuk menentukan kelas dari suara tersebut. Setelah suara terklasifikasi, *transmitter* inframerah akan mengirim sinyal inframerah sesuai dengan kelas dari suara untuk mengendalikan perangkat inframerah.

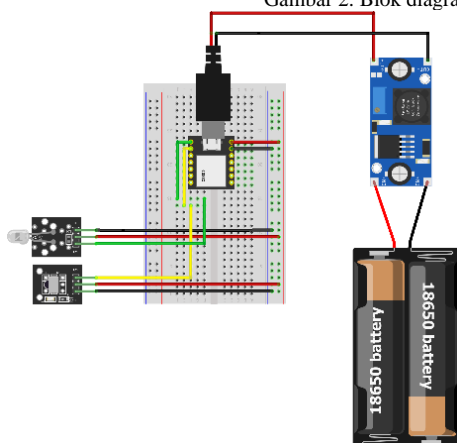
Pada arsitektur CNN, suara akan diproses oleh blok Reshape untuk mengubah suara menjadi bentuk matrix 4 dimensi. Setelah dilakukan Reshape akan diteruskan ke blok Conv2D untuk melakukan ekstraksi fitur pada data yang lalu akan diproses oleh fungsi aktivasi Relu yang berfungsi untuk memperkecil data dengan cara menghilangkan data dengan nilai dibawah 0. Setelah itu akan masuk lagi ke blok Reshape kedua dimana data akan di flatten matrix 1 dimensi. Setelah menjadi matrix 1 dimensi, data akan dimasukkan ke dalam blok

FullyConnected dimana pengklasifikasian kelas dilakukan.

Proses



Gambar 2. Blok diagram proses training dan klasifikasi (pengujian).



Gambar 3 Perancangan Keseluruhan Perangkat Keras

Setelah kelas terklasifikasi, data akan masuk kedalam blok Softmax dimana data yang tadinya mempunyai rentang $x - y$ menjadi rentang yang sudah ditentukan. Setelah Softmax dilakukan, akan lanjut kebagian terakhir yaitu output. Output yang akan didapatkan adalah kelas dengan score tertinggi.

3.3 Implementasi Sistem

Seluruh komponen kecuali baterai dan step-down LM2596 akan diletakkan kedalam kotak. *Transmitter* inframerah ditaruh pada lubang yang sudah dibuat di bagian depan kotak, *Receiver* inframerah diletakkan pada lubang yang sudah dibuat pada bagian kanan boks, baterai diletakkan kedalam *battery holder* yang diletakkan pada bagian kiri boks, dan step-down LM2596 akan diletakkan pada bagian belakang boks yang dapat dilihat pada Gambar 5.

4. PENGUJIAN DAN ANALISIS

4.1 Jarak Pengiriman dan Penerimaan Inframerah

Pada pengujian ini, diujikan jarak penerimaan receiver inframerah dan juga pengiriman transmitter

inframerah. Pengujian ini dilakukan tanpa adanya halangan antara remote ataupun perangkat yang diujikan. Dilakukan pengujian menggunakan dua remote berbeda dan dua perangkat berbeda. Remote yang digunakan adalah remote TV merek Samsung dan remote kipas angin merek Robot. Perangkat yang diujikan untuk dikendalikan adalah TV merek Samsung dan kipas angin merek Robot. Pengujian dilakukan dengan jarak 30cm, 60cm, 100cm. Hasil pengujian pengiriman inframerah dapat dilihat pada Tabel 2. Hasil penerimaan inframerah dapat dilihat pada Tabel 3.

Tabel 2. Pengujian Pengiriman Inframerah

Jarak	Samsung	Robot
30 cm	5	5
60 cm	5	3
100 cm	5	0

Tabel 3. Pengujian Penerimaan Inframerah

Jarak	Samsung	Robot
30 cm	5	5
60 cm	5	5
100 cm	5	5

Dari hasil pengujian yang telah dilakukan, dapat dilihat bahwa jarak optimal untuk mengirim inframerah berada pada jarak 30 cm dan jarak optimal untuk menerima sinyal inframerah berada pada jarak 30 cm - 100 cm. Perlu diperhatikan juga pada Tabel 1, walaupun jarak pengiriman sama, tetapi perangkat Robot jauh lebih sedikit menerima saat dikirimkan sinyal inframerah. Ini menandakan bahwa perangkat yang dikendalikan juga berpengaruh terhadap jarak pengiriman sinyal inframerah.

4.2 Pengaruh Step Terhadap Model CNN

Disini akan dibahas hasil dari pengujian jumlah *step* terhadap model CNN. Ada dua hal yang akan dianalisis yaitu pengaruh *step* terhadap akurasi model CNN.

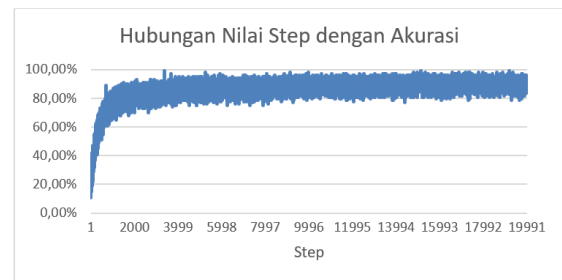


Gambar 4. Alur Program Utama



Gambar 5. Implementasi Sistem

Dapat dilihat pada Gambar 6. bahwa semakin tinggi nilai step, maka akan semakin bagus akurasi. Nilai akhir akurasi pada step ke-20000 adalah 93%.



Gambar 6. Hubungan Step dengan Akurasi

4.3 Akurasi Model CNN terhadap Suara Pengguna

Akurasi model CNN terhadap suara pengguna disajikan pada Tabel 4.

Kelas	Jumlah Sesuai	Persentase Akurasi (sesuai/200)
Nyala	154	77%
Mati	168	84%
Naik	149	77%
Turun	118	59%
Background	197	98.5%

Dapat dilihat dari Tabel 4 bahwa kelas background merupakan kelas dengan akurasi paling tinggi dan kelas Turun merupakan kelas dengan akurasi paling rendah. Terdapat 5 kelas dimana akurasi kelas Nyala adalah 77%, kelas Mati adalah 84%, kelas Naik adalah 77%, kelas Turun adalah 59%, dan kelas background adalah 98,5%. Rata – rata dari 4 kelas utama adalah 74,25% pada 200 kali pengujian di tiap kelas dengan total 800 pengujian pada 4 kelas utama.

4.4 Jarak Mikrofon Dengan Pengguna

Hasil evaluasi pengujian sistem ditinjau dari jarak ditunjukkan pada Tabel 5.

Jarak	Kelas Utama	Semua Kelas (Utama + background)
15 cm	80%	83%
30 cm	78%	82%
45 cm	72%	74%
60 cm	63%	66%

Dapat dilihat pada Tabel 5, bahwa jarak yang memiliki akurasi paling tinggi untuk klasifikasi kelas utama ada di jarak 15 cm. Tetapi, perbedaan akurasi antar jarak 15 cm dengan jarak 30 cm relatif kecil, jadi jarak optimal untuk peletakan sistem adalah 15 cm – 30 cm dari kepala pengguna.

4.5 Waktu Komputasi

Pada pengujian waktu komputasi akan diukur berapa lama waktu yang dibutuhkan sistem untuk mengklasifikasikan suara pengguna. dapat dilihat pada Tabel 5, bahwa waktu komputasi yang diperlukan untuk pengklasifikasian suara pengguna

ada diantara range 0.104 detik - 0.108 detik. Rata rata waktu komputasi dari pengujian pada Tabel 6 adalah 0,1054 detik.

Tabel 6. Waktu Komputasi

No	Waktu(s)
1	0,104
2	0,104
3	0,106
4	0,105
5	0,108
6	0,107
7	0,104
8	0,105
9	0,105
10	0,106

5. KESIMPULAN

Kesimpulan yang didapatkan dari penelitian ini adalah,

1. Dari uraian di atas, dapat disimpulkan bahwa akurasi model CNN dalam melakukan speech recognition pada pengguna cukup rendah, yaitu sekitar 74-77%. Namun, hasil tersebut dinilai cukup untuk sistem melakukan speech recognition. Kemudian, kelas dengan akurasi tertinggi adalah kelas "Mati" dengan akurasi 84%, sementara kelas dengan akurasi terendah adalah kelas "Turun" dengan akurasi 59%.
2. Dari uraian di atas, dapat disimpulkan bahwa ada hubungan yang positif antara nilai step pada fase training dengan akurasi model CNN. Semakin tinggi nilai step, maka semakin tinggi pula akurasi yang didapatkan. Pada pengujian dengan 20000 step, akurasi yang didapatkan adalah 93% yang terbilang cukup memuaskan.
3. Dari uraian di atas, dapat disimpulkan bahwa rata-rata waktu yang diperlukan sistem untuk melakukan speech recognition adalah 0,105 detik. Ini dinilai cukup memuaskan karena sistem dapat menjalankan speech recognition dengan cepat.
4. Dari uraian di atas, dapat disimpulkan bahwa jarak optimal antara pengguna dengan mikrofon berada di jarak antara 15 cm - 30 cm. Pada jarak tersebut, akurasi yang didapatkan adalah 80% untuk kelas utama pada jarak 15 cm dan 78% untuk kelas utama pada jarak 30 cm. Namun, pada jarak 45 cm dan 60 cm, hasil akurasi kurang bagus. Hal ini dinilai kurang memuaskan karena peneliti menginginkan sistem bisa mendeteksi suara dari lebih jauh.
5. Dari uraian di atas, dapat disimpulkan bahwa jarak optimal antara transmitter inframerah dengan perangkat yang dikendalikan adalah 30 cm. Hal ini dinilai kurang memuaskan karena peneliti menginginkan sistem bisa mengirim sinyal inframerah pada jarak 100 cm ke atas. Penelitian ini juga menunjukkan bahwa walaupun transmitter infrared bisa mengirimkan sinyal dengan baik, perangkat yang dikendalikan tidak selalu bisa menerimanya

dengan baik. Ini terlihat jelas dari fakta bahwa TV merek Samsung dapat dengan sukses menerima semua sinyal inframerah yang dikirimkan pada semua jarak yang diujikan, sementara kipas angin merek Robot hanya bisa menerima dengan baik pada jarak 30 cm.

Berdasarkan hasil pengujian, keefektifan sistem yang diusulkan cukup representatif diterapkan pada pengguna dengan keterbatasan (disabilitas) sesuai dengan tujuan awal penelitian ini.

DAFTAR PUSTAKA

- ABDEL-HAMID, O., MOHAMED, A. R., JIANG, H., DENG, L., PENN, G., & YU, D., 2014. Convolutional neural networks for speech recognition. *IEEE Transactions on Audio, Speech and Language Processing*, 22(10), 1533–1545.
- ABDUL QAYYUM, A.B., AREFEEN, A. AND SHAHNAZ, C., 2019. Convolutional Neural Network (CNN) based speech-emotion recognition. *2019 IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON)*, pp.122–125.
- AJAEGBU, C., ADETUNJI, O., NWAOGA, N., JULIANA, N., 2020. A Speech Activated Control System for Infrared Appliances. *International Information and Engineering Technology Association (IETA)*, 53:1, pp. 103-110.
- AOUANI, H., & BEN AYED, Y., 2018. Emotion recognition in speech using MFCC with SVM, DSVM and auto-encoder. *2018 4th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, pp. 1–5.
- BAJPAI, S., & RADHA, D., 2019. Smart Phone as a Controlling Device for Smart Home using Speech Recognition. *2019 International Conference on Communication and Signal Processing (ICCSP)*, 0701–0705. <https://doi.org/10.1109/ICCSP.2019.8697923>
- BADSHAH, A.M., AHMAD, J., RAHIM, N. AND BAIK, S.W., 2017. Speech emotion recognition from spectrograms with deep convolutional neural network. *2017 International Conference on Platform Technology and Service (PlatCon)*, pp.1–5.
- CURILEM, M., CANARIO, J. P., FRANCO, L., & RIOS, R. A., 2018. Using CNN To Classify Spectrograms of Seismic Events From Llaima Volcano (Chile). *2018 International Joint Conference on Neural Networks (IJCNN)*, 1–8.
- KHAN, A. S., AHMAD, Z., ABDULLAH, J., & AHMAD, F., 2021. A Spectrogram Image-Based Network Anomaly Detection System Using Deep Convolutional Neural Network. *IEEE Access*, 9, 87079–87093.

- KUMAR, C., UR REHMAN, F., KUMAR, S., MEHMOOD, A., & SHABIR, G., 2018. Analysis of MFCC and BFCC in a speaker identification system. 2018 International Conference on Computing, Mathematics and Engineering Technologies (ICoMET), 1–5.
- MENG, H., YAN, T., YUAN, F., & WEI, H., 2019. Speech Emotion Recognition From 3D Log-Mel Spectrograms With Deep Learning Network. IEEE Access, 7, 125868–125881. <https://doi.org/10.1109/ACCESS.2019.2938007>
- MTSHALI, P. AND KHUBISA, F., 2019. A smart home appliance control system for physically disabled people. 2019 Conference on Information Communications Technology and Society (ICTAS), pp.1–5.
- MUZAMAL, J. H., ASGHAR, M., KWONG, A., & RAZA, U. A., 2021. Microcontroller Based Intelligent Chinese-Speech Keywords Detector by Transferring the Mid-level Features of Deep Speech. 2021 International Conference on Communication Technologies (ComTech), 33–38.
- NASSIF, A. B., SHAHIN, I., ATTILI, I., AZZEH, M., & SHAALAN, K., 2019. Speech Recognition Using Deep Neural Networks: A Systematic Review. IEEE Access, 7, 19143–19165.
- PALAZ, D., MAGIMAI-DOSS, M. AND COLLOBERT, R., 2015. Convolutional Neural Networks-based continuous speech recognition using raw speech signal. 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp.4295–4299.
- PRASOMPHAN, S., 2015. Improvement of speech emotion recognition with neural network classifier by using speech spectrogram. 2015 International Conference on Systems, Signals and Image Processing (IWSSIP), pp.73–76.
- RAMLEE, R.A., OTHMAN, M.A., LEONG, M.H., ISMAIL, M.M. AND RANJIT, S.S.S., 2013. Smart Home System using Android application. 2013 International Conference of Information and Communication Technology (ICoICT), pp.277–280.
- RAMLEE, R.A., TANG, D.H. AND ISMAIL, M.M., 2012. Smart Home system for disabled people via Wireless bluetooth. 2012 International Conference on System Engineering and Technology (ICSET), pp.1–4.
- WARDEN, P., & SITUNAYAKE, D. (N.D.). TinyML Machine Learning with TensorFlow Lite on Arduino and Ultra-Low-Power Microcontrollers PREVIEW OF FIRST SIX CHAPTERS Buy the full book at tinymlbook.com.
- ZAHID, L., MAQSOOD, M., DURRANI, M. Y., BAKHTYAR, M., BABER, J., JAMAL, H., MEHMOOD, I., & SONG, O.-Y., 2020. A Spectrogram-Based Deep Feature Assisted Computer-Aided Diagnostic System for Parkinson’s Disease. IEEE Access, 8, 35482–35495.
- ZOU, Z., WANG, Q., QIN, T., WANG, Q., ZOU, B., ZHOU, M. Speed Recognition System of Smart Home. 2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Chengdu, China, 2019, pp. 2464-2468.

Halaman ini sengaja dikosongkan