

PENERAPAN SISTEM SPEAKER PINTAR YANG MENDUKUNG GOOGLE ASSISTANT

Rizqy Seto Atmaji¹, Rizki Ardianto Priramadhi², Wahmisari Priharti³

^{1,2,3} Universitas Telkom, Bandung

Tanyarizqyseto@student.telkomuniversity.ac.id¹, rizkia@telkomuniversity.ac.id²,
wpriharti@telkomuniversity.ac.id³

Abstrak

Speech to Text merupakan sistem yang mengubah suara menjadi teks. Sistem ini terdiri dari dua komponen, komponen pertama memproses sinyal akustik yang ditangkap oleh mikrofon dan komponen kedua untuk menginterpretasikan sinyal diproses, kemudian memetakan sinyal ke dalam kata-kata.

Dalam Tugas Akhir ini penulis membuat sistem *Speaker* pintar didalamnya terdapat sistem *Speech to Text*, yang berguna untuk memberikan kemudahan khususnya bagi orang berkebutuhan khusus dan lansia dalam mengendalikan peralatan elektronik dengan hanya menggunakan suara. *Speaker* pintar berfungsi untuk mengubah perintah suara menjadi teks yang akan disimpan pada *cloud* berbasis *Internet of Things* (IoT). Pengolahan suara akan diproses dengan menggunakan Google API.

Sistem yang dibangun dapat digunakan untuk mendeteksi suara pada jarak maksimal 2.1 meter dari *user* untuk perintah menghidupkan/mematikan lampu, mengontrol warna lampu dan intensitasnya. Jarak dengan akurasi paling bagus dari pengujian yakni pada jarak 50 cm dengan akurasi 91%. Dari pengujian yang didapat disimpulkan bahwa semakin jauh jarak maka semakin keras juga intensitas suara yang diperlukan dan respon waktu juga semakin lama. Daya yang dibutuhkan pada sistem ini rata rata 3.08 watt.

Kata Kunci: *Smart Home, Speech to Text, Internet of Things, Sistem.*

Abstract

Speech to Text is a system that converts voice into text. This system consists of two components, the first component processes the acoustic signal captured by the microphone and the second component interprets the processed signal, then maps the signal into words.

In this final project the author makes a smart *Speaker* system in which there is a *Speech to Text* system, which is useful for providing convenience, especially for people with special needs and the elderly in controlling electronic equipment using only voice. The smart *Speaker* functions to convert voice commands into text that will be stored in the *Internet of Things* (IoT)-based cloud. Voice processing will be processed using Google API.

The system built can be used to detect sound at a maximum distance of 2.1 meters from the user for commands to turn on/off the lights, control the color of the lights and their intensity. The distance with the best accuracy from the test is at a distance of 50 cm with an accuracy of 91%. From the test, it can be concluded that the farther the distance, the louder the intensity of the sound required and the response time will also be longer. The power required in this system is an average of 3.08 watts.

Keywords: *Smart Home, Speech to Text, Internet of Things, System.*

1. Pendahuluan

Perkembangan teknologi *smart home* yang sangat pesat, *smart home* ada dikarenakan ada masalah pada orang berkebutuhan khusus dan lansia yang kesulitan bergerak untuk mengendalikan perangkat elektronik rumah tangga selain itu orang kesulitan dalam mengerjakan pekerjaan dalam waktu yang sama.

Dengan adanya *smart home* maka orang berkebutuhan khusus dan lansia dapat memudahkan dalam mengendalikan peralatan elektronik rumah tangga, yang dapat dikontrol dari jarak jauh. *smart home* membuat seluruh peralatan rumah tangga terhubung dengan sistem otomasi canggih[1]. Tujuan dibuatnya *smart home* adalah untuk meningkatkan

keamanan,kenyamanan, dan kemudahan menggunakan perangkat elektronik rumah tangga. Dalam pengendalian *smart home* biasanya dengan perintah suara pada *smartphone*[2].

Pada penelitian sebelumnya terdapat beberapa kekurangan yang sudah dijelaskan diatas, maka dari itu akan dibuat *speaker* pintar sebagai solusinya, peran *speaker* pintar ini bisa menggantikan perintah suara *smartphone*. Dalam penelitian pengolahan suara menggunakan Google Cloud *speech* API. Google Cloud *speech* API merupakan salah satu layanan *machine learning* Google dimana Google Cloud *speech* API juga termasuk dalam *Artificial Intellegent*[3]. Penelitian ini berbasis IoT yaitu menggunakan *platform* IoT *Firebase*.

Berdasarkan pada uraian diatas, sistem *speaker* pintar ini dibuat untuk membantu dan mempermudah *user* dalam mengendalikan perangkat listrik terkhususnya pengendalian lampu antara lain pengendalian warna lampu,intensitas lampu dan hidup atau mati lampu dengan menggunakan perintah suara.

2. Dasar Teori

2.1 Internet of Things

Internet of Things adalah pengembangan teknologi dalam konsep transfer data. Dengan adanya IoT, proses transfer data dapat berjalan tanpa adanya campur tangan manusia. Setiap perangkat sistem dapat terintegrasi satu sama lain berdasarkan program dan sistem komunikasi yang dibangun oleh manusia[4]. Ketika sistem IoT sudah berjalan, manusia hanya berperan sebagai pengatur dan pengawas sistem. *Internet of Things* memiliki banyak kegunaan, contohnya adalah untuk *monitoring* dan *controlling* suatu sistem hanya dengan menggunakan jaringan internet. Kemampuan-kemampuan yang dimiliki IoT sangat berdampak bagi konsumen, baik perusahaan-perusahaan maupun masyarakat secara keseluruhan khususnya dalam peningkatan layanan kehidupan, seperti *smart home*, *smart cities*, meningkatkan kualitas kesehatan, pendidikan, dan produktivitas[5].

2.2 Firebase

Firebase merupakan platform untuk aplikasi realtime. Ketika data berubah, maka aplikasi yang terhubung dengan *firebase* akan meng-*update* secara langsung melalui setiap *device* (perangkat) baik *website* ataupun *mobile*[6]. *Firebase* mempunyai *library* (pustaka)

yang lengkap untuk sebagian besar *platformweb* dan *mobile* dan dapat digabungkan dengan berbagai *framework* lain seperti *node*, *java*, *javascript*, dan lain-lain. *Application Programming Interface* (API) untuk menyimpan dan sinkronisasi data akan disimpan sebagai *bit* dalam bentuk *JSON* (*JavaScript Object Notation*) pada *cloud* dan akan disinkronisasi secara *realtime*.

2.3 Speech Recognition

Speech Recognition (pengenalan kalimat atau kata) merupakan ilmu komputer dan teknik elektronika merupakan sebuah sistem yang mengubah kalimat suara menjadi kode-kode digital yang berfungsi sebagai perintah untuk melakukan sesuatu pada sistem. Konsep teknologi yang dikembangkan dalam *Speech Recognition* biasanya disebut *Voice Command System*[7].

Dalam penelitian ini sistem *Speech Recognition* yang dipakai yaitu *Google Cloud Speech API* merupakan suatu produk AI dan *Machine Learning* milik google yang dapat mengkonversi audio-teks yang biasa disebut *Google Cloud Speech-to-Text API* dan *Google Cloud Text to Speech API* yang dapat digunakan untuk konversi teks-audio[8].

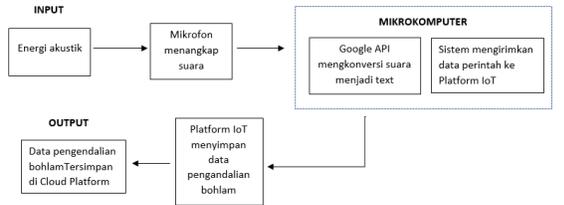
2.4 Google Speech API

Google Speech API adalah sebuah kerangka kerja yang dibuat oleh Google untuk mengenali suara, dirubah menjadi string (teks) dan dimasukkan ke dalam halaman pencarian Google sehingga akan menampilkan hasil pencarian berdasarkan masukan suara tersebut. Pengenalan suara dilakukan pada server Google menggunakan algoritma Hidden Markov Model (HMM)[9]. Dengan kata lain input suara yang diterima oleh perangkat akan dikirimkan ke server Google, yang selanjutnya server Google melakukan pengenalan dan mengubahnya menjadi teks menggunakan algoritma HMM. Hasil konversi berupa teks[9].

Google Speech adalah suatu API yang disediakan oleh Google untuk mengidentifikasi suara dengan menggunakan cara digitalisasi kata dan mencocokkan sinyal digital tersebut dengan suatu pola yang tersimpan di dalam database Google. Teknologi *Speech Recognition* merupakan teknologi pengenalan kata yang memanfaatkan sinyal suara manusia sebagai masukan untuk kemudian dikenali oleh sistem komputer[9].

3. Perancangan Sistem

3.1 Diagram Blok Sistem



Gambar 3. 1 Diagram blok sistem

Berdasarkan Gambar, Masukan dari sistem adalah perintah suara dari *user*. Perintah suara akan diterima oleh *microphone* yang sudah terintegrasi pada mikrokomputer. Mikrokomputer yang digunakan pada sistem adalah Raspberry Pi yang merupakan otak utama dari sistem yang berperan untuk menerima data suara dari *microphone* dan melakukan proses pengolahan suara menjadi teks menggunakan Google API. Hasil dari pengolahan suara berupa status pengendalian bohlam akan dikirim ke internet, dimana *cloud* yang digunakan adalah *Firebase*.

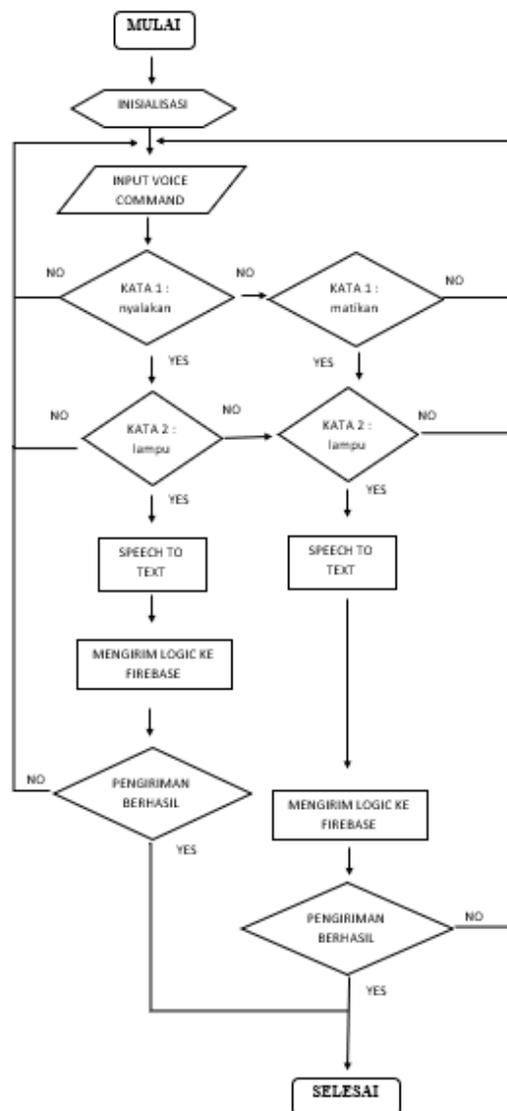
3.2 Desain Perangkat Keras



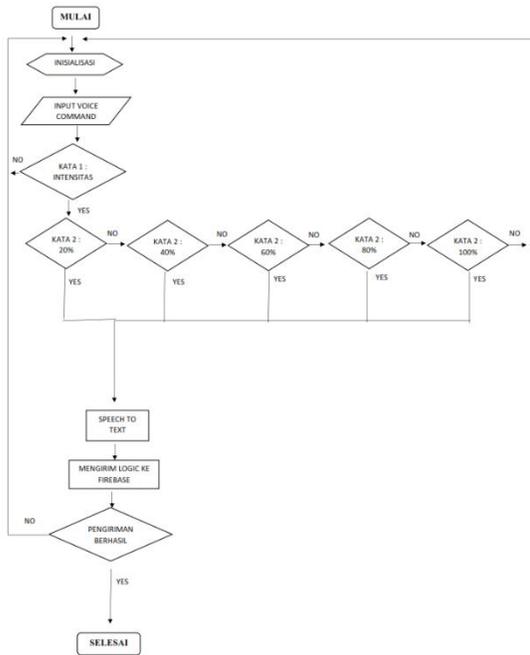
Gambar 3. 2 Desain perangkat keras

Berdasarkan Gambar, Desain perancangan perangkat keras *power supply* dari raspi dan speaker terhubung dengan Power bank 1000 Mah. Mic dan speaker terhubung ke *sound card*, mic sebagai *input* suara dan speaker sebagai *Output* suara selanjutnya *sound card* akan terhubung dengan raspberry pi melalui USB.

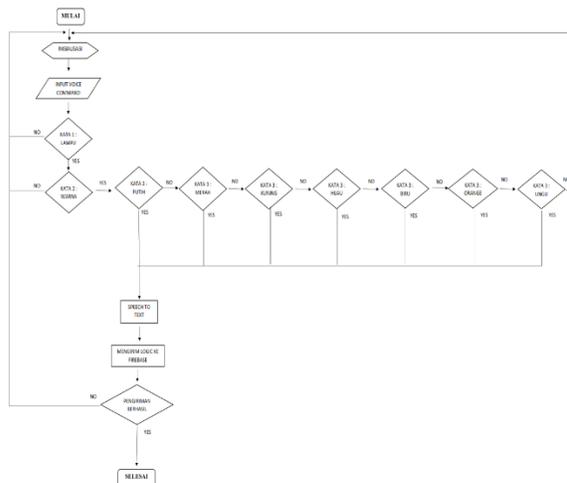
3.3 Diagram Alir Sistem



Gambar 3. 3 Diagram Alir Sistem nyala mati lampu



Gambar 3. 4 Diagram Alir Sistem Intensitas



Gambar 3. 5 Diagram Alir Sistem Warna Lampu

Berikut alur sistem pada Pengendalian Bohlam.

- Sistem dimulai dari proses inialisasi data berupa *input* dan *output*, mengakses *library* atau *tools* yang digunakan untuk mendukung sistem.
- Selanjutnya, *microphone* akan menangkap perintah suara sebagai *input*. *Microphone* sudah terintegrasi pada Raspberry Pi.
- Suara yang diterima selanjutnya diproses untuk diubah dari sinyal audio menjadi teks. Proses ini disebut *speech to text*. *Speech to text* merupakan

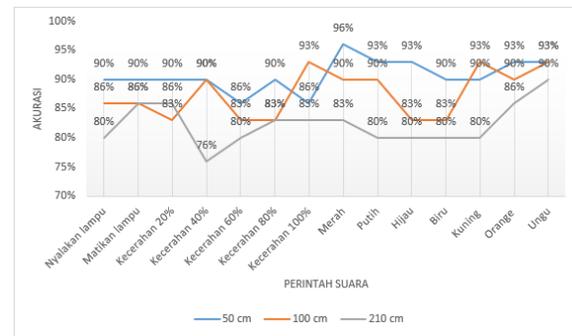
fitur yang terdapat pada Google API yang sudah terintegrasi pada Raspberry Pi.

- Hasil dari proses *speech to text* akan diidentifikasi tiap karakter yang terkandung di dalamnya dan akan disesuaikan dengan *keyword* yang sudah ditetapkan untuk mengendalikan bohlam.
- Hasil akhir dari proses pengolahan suara adalah berupa *logic*. *Logic* yang digunakan untuk mengendalikan bohlam sesuai perintah.
- Akhirnya, *logic* akan dikirim ke *Firestore*.

4. Hasil Pengujian dan Analisis

Pada bab ini akan dilakukan pengujian berdasarkan perancangan sistem yang telah dibuat. Pengujian dilakukan agar mengetahui apakah sistem sudah dapat berjalan sesuai perencanaan.

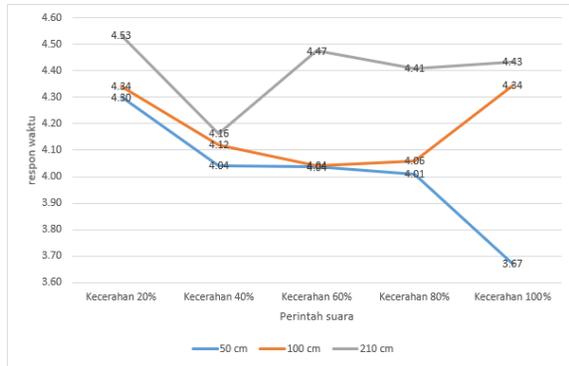
4.1 Pengujian terhadap keakurasian sistem perintah suara dari jarak yang berbeda



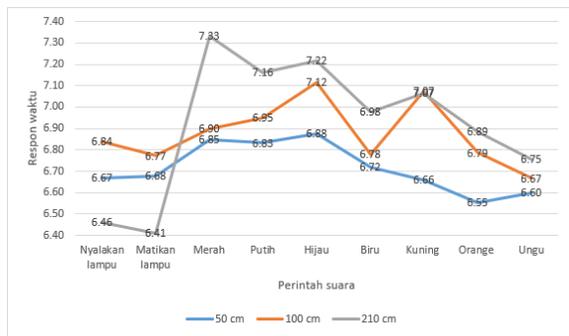
Grafik 4. 1 Pengujian terhadap keakurasian sistem perintah suara dari jarak yang berbeda

Berdasarkan grafik, didapatkan rata rata keberhasilan dari jarak 50 cm, 100 cm, dan 210 cm yaitu 91%, 88%, dan 82% maka dapat disimpulkan bahwa semakin jauh *Speaker* dari penguji maka semakin sedikit juga presentase keberhasilan. Selain itu, kegagalan juga dapat disebabkan juga oleh koneksi internet yang buruk ketika pengujian dilakukan. Karena berbasis *Internet of Things*, maka internet sangat berpengaruh dalam jalannya fungsi sistem yang dibangun. Ketika koneksi internet lemah atau *lost connection* maka tingkat akurasi keberhasilan sistem dalam memproses perintah juga akan menjadi lemah. Namun, sebaliknya, ketika koneksi internet bagus, maka tingkat akurasi keberhasilan sistem dalam memproses perintah juga akan semakin bagus. Selain itu kegagalan juga dikarenakan gangguan suara dari luar.

4.2 Pengujian terhadap respon waktu sistem pendeteksi perintah dengan jarak yang berbeda



Grafik 4. 2 Pengujian perintah kecerahan terhadap respon waktu sistem pendeteksi perintah dengan jarak yang berbeda

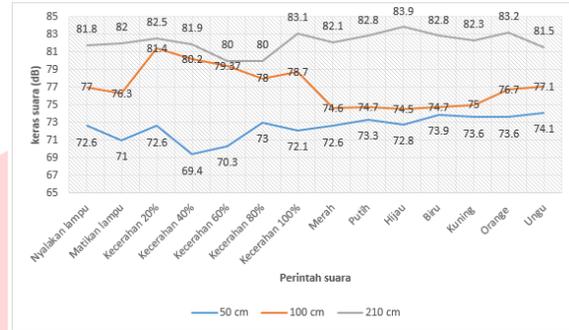


Grafik 4. 3 Pengujian perintah warna dan hidup mati lampu terhadap respon waktu sistem pendeteksi perintah dengan jarak yang berbeda

Berdasarkan grafik, terdapat perbedaan waktu antara grafik respon waktu kecerahan dan respon waktu nyala mati dan warna. Perbedaan waktu tersebut dikarenakan pada *firebase*, pada perintah mati, hidup dan perubahan warna perintah data utama yang dirubah sedangkan untuk pengaturan kecerahan yang dirubah sedangkan pada perintah kecerahan yang dirubah adalah sub data perintah kecerahan analisis yang didapat bahwa sub data memproses perubahan perintah lebih cepat dibandingkan dengan perintah data utama. Pada pengujian jarak 50 cm, 100 cm, dan 210 cm terhadap respon waktu, jarak juga berpengaruh terhadap respon waktu jadi semakin jauh jarak maka semakin lama juga respon waktunya, sedangkan kecepatan internet juga berpengaruh terhadap respon waktu pada saat pengujian jarak 210 cm kecepatan internet unduh dan unggah, 36.2 Mbps dan 35.9 Mbps.

Pada pengujian jarak 100 cm kecepatan internet 43.5 Mbps unduh dan 27.7 Mbps unggah. Pengujian pada jarak 50 cm kecepatan unduh 52.6 Mbps dan unggah 36.6 Mbps. Pada pengujian ini menggunakan aplikasi *speedtest* pada ios.

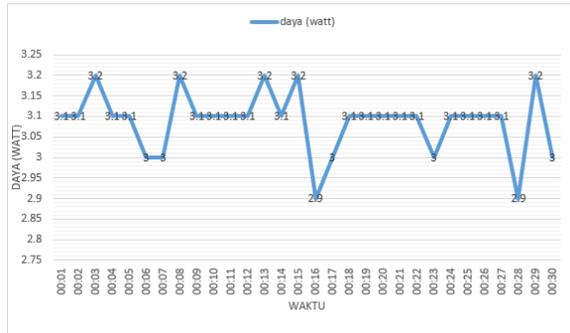
4.3 Pengujian terhadap intensitas suara perintah dari jarak yang berbeda



Grafik 4. 4 Pengujian terhadap intensitas suara perintah dari jarak yang berbeda

Berdasarkan grafik, maka dapat dihasilkan rata rata intensitas suara dari setiap jarak berbeda beda yaitu pada jarak 50 cm didapatkan intensitas suara 72.5 dB, jarak 100 cm intensitas suaranya 77 Db, dan pada jarak 210 cm intensitas suaranya 82.1 Db. Maka dapat disimpulkan bahwa semakin jauh jaraknya, semakin besar pula intensitas suara yang dibutuhkan, agar *Speaker* dapat mendeteksi perintah. Dalam pengujian ini letak db meter jaraknya 30 cm dari penguji. Dari pengujian ini didapatkan pada jarak 50 cm intensitas suara terendah 62 db dan intensitas tertinggi 83 db, sedangkan pada jarak 100 cm intensitas suara terendah 67 db dan intensitas tertinggi 94 db, dan pada jarak 210cm intensitas terendah 69 db dan intensitas tertinggi 94 db. Berdasarkan pengujian yang telah dilakukan bahwa dapat disimpulkan rata rata suara yang dihasilkan termasuk skala kuat.

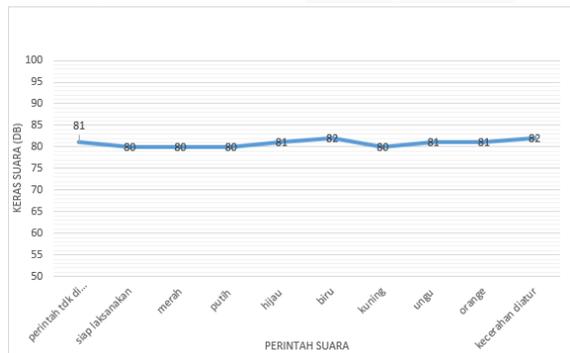
4.4 Pengujian daya yang dibutuhkan sistem



Grafik 4. 5 Pengujian daya yang dibutuhkan sistem

Berdasarkan grafik, daya rata rata yang dibutuhkan sistem 3.08 watt. Pada pengujian ini daya tertinggi pada 3.2 watt dan daya terendah 2.9 watt. Perbedaan daya tersebut dikarenakan pada saat *Speaker* menghasilkan suara maka watt yang diperlukan bertambah maka dari itu terjadi perbedaan watt yang dibutuhkan. Pengujian ini menggunakan alat watt meter. Berdasarkan pengujian yang telah dilakukan dapat disimpulkan bahwa sistem ini membutuhkan daya yang rendah maka dari itu sistem ini termasuk sistem hemat daya.

4.5 Pengujian intensitas suara respon sistem



Grafik 4. 6 Pengujian intensitas suara respon sistem

Berdasarkan grafik, maka dapat disimpulkan bahwa intensitas suara rata rata respon sistem yakni 81 db pada respon. Intensitas suara diperoleh dari pengukuran respon suara sistem diukur dengan aplikasi dB meter dengan jarak antara *Speaker* dan alat penguji 10 cm. untuk seharusnya respon suara sistem harusnya sama akan tetapi pada pengujian ini ada perbedaan dikarenakan ada noise dari luar selama pengujian dilakukan. Pengujian ini dilakukan di ruang

kamar ukuran 3 x 3 meter maka dari itu dapat disimpulkan bahwa rata rata intensitas suara yang dihasilkan dapat terdengar dengan jelas.

5. Kesimpulan

Berdasarkan perancangan, pengujian dan analisis pada Tugas Akhir diperoleh kesimpulan sebagai berikut.

1. Pada Tugas Akhir ini berhasil dibangun sistem pendeteksi perintah suara dengan menggunakan *google speech API*, sistem dapat mendeteksi perintah suara dengan jarak maksimal 2.1 meter dalam ruangan.
2. Pada Tugas Akhir ini berhasil mengintegrasikan sistem pendeteksi perintah suara ke dalam sistem *Internet of Things*, dimana dalam proses mengirim data suara sampai *server firebase*, sistem berhasil menjalankan proses dari semua perintah dan diperoleh akurasi rata rata pada jarak 50,100, dan 210 cm adalah 91%,88% dan 82%.
3. Diperoleh hasil untuk jarak ideal sumber suara terhadap *microphone* terdapat pada jarak 50 cm, karena memiliki rata-rata akurasi yang paling bagus yaitu 91%.
4. Pengujian intensitas suara berbasis *Internet of Things* dilakukan pengujian di 3 jarak, yaitu 50 cm, 100 cm, dan 210 cm. Berdasarkan pengujian yang dilakukan, diperoleh kesimpulan bahwa semakin jauh *Speaker* maka semakin keras juga suara yang dibutuhkan.
5. Pengujian untuk respon waktu semakin jauh *Speaker* semakin lama pula respon waktu selain itu mempengaruhi respon waktu yaitu kecepatan internet yang sedang dipakai.
6. Pada sistem ini rata rata daya yang dibutuhkan 3.08 watt.
7. Dari sistem ini di dapat intensitas suara respon sistem rata rata 80.3 db.

6. Saran

Berdasarkan perancangan, implementasi, pengujian, analisis, dan kesimpulan yang sudah ada, berikut beberapa saran dari penulis untuk kajian atau penelitian lanjutan serta *practical implication* mengenai sistem yang dibangun pada Tugas Akhir ini.

1. Untuk penelitian selanjutnya diharapkan *feedback Speaker* tidak hanya suara dari google,

melainkan dapat menggunakan suara *public figure*.

2. Untuk penelitian selanjutnya, diharapkan untuk mengembangkan sistem ke tahap yang lebih lanjut yaitu bukan hanya mengenali perintah suara, namun lebih *next level* lagi yaitu untuk mengenali suara *Speaker* yang memberi perintah ke sistem, sehingga sistem akan lebih aman dan kompleks.

Referensi

- [1] T. Malche and P. Maheshwary, "Internet of Things (IoT) for building smart home system," *Proc. Int. Conf. IoT Soc. Mobile, Anal. Cloud, I-SMAC 2017*, no. February 2019, pp. 65–70, 2017, doi: 10.1109/I-SMAC.2017.8058258.
- [2] A. Purwanto and S. Lutfi, "PENGENDALIAN LAMPU RUMAH BERBASIS GOOGLE ASISSTANT MELALUI SMARTPHONE MENGGUNAKAN NodeMCU-12E ESP8266 DI NUKE KOMPUTER SERVICE," *J. Himsya Tech*, vol. 20, no. 2, pp. 1–6, 2019.
- [3] A. Ravulavaru, *AI Services Quick Start Guide*. Packt Publishing Ltd, 2018.
- [4] A. . Rozik, A. S. Tolba, and M. A. El-Dosuky, "Design and Implementation of the Sense Egypt Platform for Real-Time Analysis of IoT Data Streams," 2016.
- [5] L. Atzori, G. Morabito, and A. Iera, *Understanding the Internet of Things (IoT)*. GSMA, 2014.
- [6] E. A. W. Sanad, "Pemanfaatan Realtime Database di Platform Firebase Pada Aplikasi E-Tourism Kabupaten Nabire," *J. Penelit. Enj.*, vol. 22, no. 1, pp. 20–26, 2019, doi: 10.25042/jpe.052018.04.
- [7] G. Team, "Professional Sound System Specialist," *PT. G`oshen Swara Indonesia*, 2016. .
- [8] G. Cloud, "AI & Machine Learning Products," *Google*, 2008. .
- [9] A. Akbar, A. Y. Husodo, and A. Zubaidi, "Implementasi Google Speech API Pada Aplikasi Koreksi Hafalan Al-Quran Berbasis Android," *J. Teknol. Informasi, Komputer, dan Apl.*, vol. 1, no. 1, pp. 1–8, 2019.