Hello World
Jurnal Ilmu Komputer
https://jurnal.ilmubersama.com/index.php/hello_world

Metode *Synthesizer Concatenation* Algoritma *Time Domain Psola* pada Aplikasi Konversi File Teks Ke Audio

Nurhidayah, Tasliyah Haramaini

Fakultas Teknik, Teknik Informatika, Universitas Islam Sumatera Utara, Medan, Indonesia

INFORMASIARTIKEL

Diterima Redaksi: 18 Juli 2024
Revisi Akhir: 19 Agustus 2024
Diterbitkan *Online*: 24 Agustus 2024

KATA KUNCI

Text-to-Speech
Synthesizer
Time Domain PSOLA

KORESPONDENSI (*)

Phone: +62 898-0671-066
E-mail: nurhid430700@gmail.com

A B S T R A K

Teknologi *text-to-speech* (TTS) telah berkembang pesat untuk berbagai bahasa, termasuk dalam konteks masyarakat Indonesia yang memiliki tingkat literasi yang rendah. Ini merupakan solusi potensial bagi individu yang menghadapi kesulitan dengan literasi karena banyaknya teks yang harus dibaca atau ditulis. Khususnya dalam pendidikan inklusif bagi penyandang cacat jasmani, kemampuan untuk membaca secara verbal melalui TTS sangat penting dalam proses pembelajaran. Penelitian ini menggunakan metode *synthesizer concatenation* dengan pendekatan Algoritma *Time Domain PSOLA* untuk menghasilkan TTS. Data penelitian dikumpulkan melalui studi pustaka, dengan bahasa pemrograman utama yang digunakan adalah PHP. Pengujian dilakukan menggunakan blackbox testing untuk memvalidasi kinerja TTS yang dikembangkan.Teknologi TTS tidak hanya memfasilitasi aksesibilitas literasi bagi masyarakat umum, tetapi juga meningkatkan aksesibilitas pendidikan untuk individu dengan kebutuhan khusus seperti penyandang cacat jasmani.

PENDAHULUAN

Kebutuhan akan teknologi terus meningkat seiring dengan tuntutan zaman. Teknologi dikembangkan untuk mempermudah kegiatan manusia dengan fokus pada kecepatan dan keakuratan.

Text-to-Speech (TTS) adalah sistem yang mengubah teks masukan menjadi sinyal ucapan yang disintesis [1]. Sintesis ucapan adalah produksi ucapan manusia secara artifisial. Sebuah sistem komputer yang digunakan untuk tujuan ini tersebut komputer ucapan atau penyintesis ucapan, dan bisa juga diimplementasikan dalam produk perangkat lunak atau perangkat keras [2].

Teknologi ini dapat diimplementasikan baik dalam perangkat lunak maupun perangkat keras. Sebagian orang yang memiliki masalah dengan literasi salah satunya dikarenakan jumlah teks yang perlu dibaca atau ditulis [3]. Masyarakat Indonesia yang rata-rata tingkat literasinya rendah juga bukanlah pengecualian. Dalam pendidikan khusus, yaitu pendidikan bagi penyandang cacat jasmani, membaca adalah penting dalam pembelajaran. Tindakan membaca memungkinkan siswa untuk mempelajari kosa kata dan konsep dan untuk mengakses berbagai jenis bahan bacaan. Ini mengapa orang yang mampu wajib membaca buku, materi kuliah dan lain-lain teks untuk pendengaran orang tuna netra sehingga memungkinkan secara fisik tantangan untuk belajar [4]. Penderita disleksia dan orang lain yang kesulitan membaca dengan lancer dalam Bahasa kedua [5] .

https://doi.org/10.56211/helloworld.v3i3.587

Dengan adanya konversi *file* teks ke audio mempermudahkan membantu berkomunikasi bagi individu dengan ganguan bicara atau kondisi lainnya yang mempengaruhi kemampuan bicara, mendapatkan informasi tanpa harus meluangkan waktu untuk membaca secara fisik, dan lebih mudah dipahami, dan lebih menarik. Menciptakan pembelajaran yang efektif adalah penggunaan media pembelajaran [6].

Time Domain PSOLA adalah teknik untuk memodifikasi pitch sinyal suara dengan memanipulasi durasi segmen suara yang tumpang tindih, menghasilkan perubahan pitch yang diinginkan dengan menjaga sinkronisasi dan kehalusan transisi antara segmen.

Metode synthesizer concatenation dengan pendekatan Algoritma Time Domain PSOLA digunakan untuk menghasilkan suara yang alami dan mudah dimengerti. Dengan memecah suara menjadi diphone dan mengaplikasikan PSOLA, sintesis ucapan dapat terdengar lancar dan natural tanpa gangguan intonasi yang berlebihan.

TINJAUAN PUSTAKA

Aplikasi

Aplikasi merupakan teknologi yang berkembang pesat. Kemajuan teknologi membantu pengelolahan data atau informasi yang tersedia dapat berlangsung secara cepat, efesien dan akurat. Aplikasi dapat mempermudahkan proses pengerjaan atau penyelesaian suatu persoalan. Pengertian aplikasi secara umum adalah alat terapan yang difungsikan secara khusus dan terpadu sesuai kemampuannya yang dimilikinya, aplikasi merupakan suatu perangkat computer yang siap pakai bagi *user.* [7]

JavaScript

JavaScript merupakan Bahasa pemprograman yang sering digunakan dalam pengembangan web. Dengan JavaScript dapat membuat interasi dan dinamika pada halaman web, seperti validasi formulir, efek animasi, manupulasi elemen HTML, serta mengelolah dan memanipulasi data [8].

Concatenation Synthesizer

*Concatenation Synthesize r*merupakan synthesizer yang mampu memproduksi sinyal ucapan secara otomatis melalui transkripsi grafem-ke-fenom untuk kalimat yang diucapkan [9]. *Concatenation* yaitu, penggabungan-gabungan segmen-segmen bunyi yang telah direkam sebelumnya. Setiap segmen berupa *Dhipone* (gabunganduabuahfenom).Pada perekamansuara yang dilakukanbeberapa kali agar mendapatkanhasil yang akurat [10]. *Synthesizer* yang menggunakan Teknik *diphone concatenation* bekerja dengan cara menggabungkan segmen-segmen bunyi yang telah direkam sebelumnya [11].

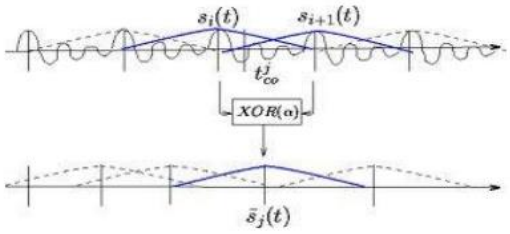
Algoritma Time Domain Psola

Time Domain PSOLA merupakan suatu teknik untuk memodifikasi *pitch* sinyal suara yang dilakukan dengan mengubahdurasi sinyal suaranya, memajang untuk mengurangi *pitch* dan memedekkan untuk meningkatkan *pitch.* Tapi dalam PSOLA gelombang suara dibagi terlebih dahulu menjadi beberapa segmen tumpang tindih kecil, kemudian segmen dipindahkan lebih dekat atau terpisah tergantung apakah akan menambah atau mengurangi *picth*. Kemudiansegmen di *overlap add*-kan, sehingga bentuk gelombang suara resultan sama dengan bentuk gelombang sebenarya. PSOLA juga dapat digunakan untuk mengubah durasi sinyal suara dengan tetap menjaga karakteristik *pitch* asli. Hal ini dilakukan dengan membuang beberapa segmen untuk mempersingkat durasi sinyal atau mengulangi beberapa segmen untuk mengingkatkan durasi sinyal.

Ada banyak jenis PSOLA, pada tugas akhir ini menggunakan *time domain* psola karena memungkiankan untuk mengubah picth sinyal suara tanpa mengubah durasinya atau sebaliknya. *Time Domain PSOLA* berfungsi untuk mensinkronkan dekomposisi *pitch* sinyal suara dengan *period*nya ketika proses *overlapping*, yang sebelumnya dilakukan *famming* terlebih dahulu [12].

 https://doi.org/10.56211/helloworld.v3i3.587

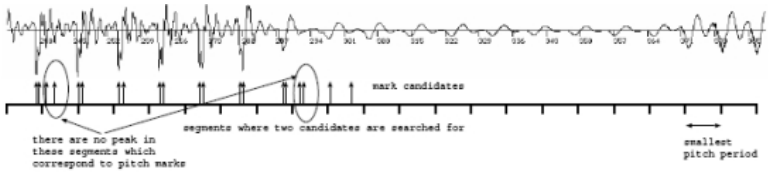
Hal-hal yang pentingdilakukan pada proses TD-PSOLA adalah:



Gambar 1. TD-PSOLA

Pitch Detection

Sinyal suara input akan terdeteksi dan dimodifikasi nilai *pitchnya*. Sinyal suara yang diproses hanya sinyal suara *voiced* saja, karena hanya sinyal suara *voiced* yang memiliki *pitch* yang bisa dihitung dan dimodifikasi. Pendeteksian *pitch* mengguanak n algoritma *pitch* mark. Algoritma ini akan mendeteksi setiap *pitch* mark dari suatu sinyal suara. *Pitch* mark merupakan beda nilai antara dua buah puncak sinyal suara. Sehingga akan diproses algoritma ini adalah *pitch* periodnya, yaitu beda nilai antar puncak sinyal suara, setelah terdeteksi semua *pitch* mark, maka langkah selanjutnya adalah membagi sinyal suara menjadi *freme-frame* [13].



Gambar 2. *Pitch Mark*

Untukpersamaan*pitch detection* yang diguanaknadalah:

C = [c(i)] = c (1)…. c(i)….c(N) (1)

D = (c(i), c(I) = | (c(I) – c(i) – *pitchPeriod* (c(i))| (2)

Keterangan:

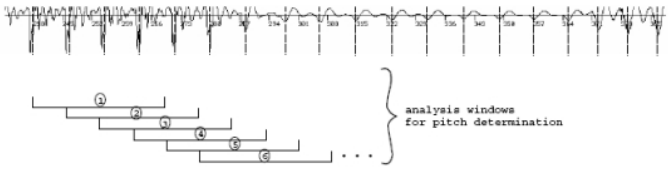
C = Frame

i = *pitch mark*

D = matriks

Framming

Panjang *frame* yang digunakan sangat mempengaruhi keberhasilan dalam analisis spectral. Di satu sisi ukuran dalam *frame* harus sepanjang mungkin untuk dapat menunjukkan resolusi frekuensi yang baik. Akan tetapi, di lain sisi ukuran *frame* juga harus cukup pendek untuk dapat menunjukkan resolusi waktu yang baik.



Gambar 3. *Framming*

Windowing

Gelombang ucapan berupa frame-frame kecil dimana setiap frame tersebut dianggap sebagai *time invariant system*. Sebuah frame ucapan x[n] didapat dari gelombang penuh signal ucapan s[n] yang dikalikan dengan fungsi “jendela” (window) w[n] dalam domain waktu:

x[n] = w[n]s[n] (3)

Proses windowing adalah suatu proses *weighting* yang berfungsi untuk mengurangi efek diskontinuitas pada ujung-ujung *frame* yang dihasilkan oleh proses *framing*. Berikut adalah blok diagram darisebuah proses *windowing* terhadap keluaran proses *framing.* Fungsi *windowing* yang digunakan adalah *window hamming*, yang mempunyai persamaan seperti berikut

$$hamming\ w(n) = 0.54 - 0.46\ \cos\ (\frac{2\times phi\times n}{(n-1)}), 0 \leq n \leq (n-1) \quad (4)$$

METODOLOGI

Penulis menganalisa dari selisih waktu yang dibutuhkan dari suara pria dan wanita. Berikut table selisih antara suara wanita dan pria.

Tabel 1. Tabel Perbandingan Suara Wanita Dan SuaraPria

Jumlah Kata	Suara Wanita (detik)	Suara Pria (detik)
10 kata	5,79 detik	4,65 detik
14 kata	7,73 detik	6,47 detik
20 kata	10,10 detik	8,93 detik
25 kata	11,65 detik	11,47 detik
35 kata	16,66 detik	16,07 detik

Kesimpulan selisih dari suara pria dan wanita adalah suara pria lebih cepat, karena suara wanita membutuhkan 1 sampai 3 detik untuk memulai membaca.

Proses Teks Ke Audio Menggunakan Metode Synthesizer Concatenation.

Synthesizer Concatenation digunakan dalam teknologi *text-to-speech* (TTS) untuk menghasilkan ucapan yang terdengar alami dengan menggabungkan segmen-segmen suara yang telah direkam sebelumnya. Berikut gambaran umum tentang bagaimana perhitungan dalam *synthesizer concatenation* dilakukan.

Misalkankitainginmenyintesiskalimat “Halo, apakabar?” dengan metode *concatenation*. Berikut langah-langkah perhitungannya secara sederhana:

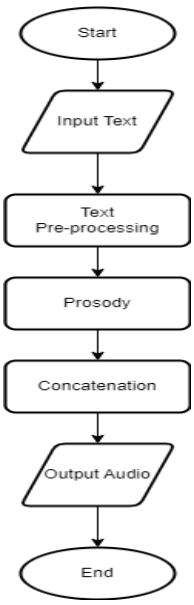
1. Tesk Pre-Processing:
   Teks dipecahkan menjadi fenom: /h/, /a/, /l/, /o/, /a/, /p/, /a/, /k/, /a/, /b/, /a/, /r/
2. Unit Selection:
   Memilih segmen-segmen suara yang sesuai dari corpus untuk setiap fenom.
3. Prosody:
   Menyesuaikan intonasi dan durasi fenom untuk memastikan kalimat “Halo, apa kabar?” diucapkan dengan nada dan tempo yang alami.
4. Concatenation dan Smoothing:
   Menggabungkan segmne-segmen yang dipilih: segmen untuk /h/ digabungkan dengan /a/, kemudian /a/, kemudian /l/, dan seterusnya.
   Melakukan smoothing unutk mengatasi perbedaan akustik segmen /h/, dan /a/, dan seterusnya.

Dengan demikian *synthesizer concatenation* menghasilkan ucapan terdengar alami dengan menggabungkan segmen-segmen suara yang sudah direkam dan diproses sedemikian rupa.

Flowchart Synthesizer Cocatenation

Berikut direfrentasikan dalam bentuk *flowchart* yang terdiri atas 3 bagian utama yaitu *text per-processing*, pengbangkitan *prodosy* dan *concatenation*. Dibawah ini *flowchart text to synthesizer system* adalah: [14].

 https://doi.org/10.56211/helloworld.v3i3.587



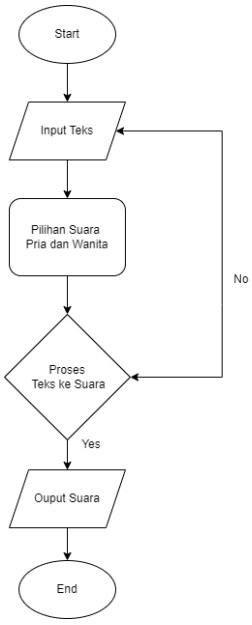
Gambar 4. *FlowchartText To Speech Synthesizer System*

Adapun tahapan proses diatas, yaitu:

1. Input Teks: Proses dimulai dengan pengguna memasukkantekssebagai input.
2. Text Pre-Processing: Teks input dikonversi menjadi diphone, yaitu gabungan dua fenom suara yang berdekatan. Jika teks mengandung akronim atau singkatan, tahap ini mengkonversinya menjadi diphone yang sesuai.
3. Prosody: Fokus pada karakteristik sinyal ucapan manusia untuk mencapai hasil ucapan yang lebih alami. Prosodi mengacu pada perubahan nilai pitch (frekuensi dasar) selama pengucapan kalimat atau sebagai fungsi waktu. Tujuannya adalah memperhalus hasil proses penggabungan diphone-diphone agar suara yang dihasilkan mendekati natural.
4. Concatenation: Penggabungan segmen-segmenbunyi yang telah direkam sebelumnya. Setiap segmen berupa diphone yang direkam beberapa kali untuk mendapatkan hasil yang akurat.
5. Output Audio: Hasil akhir dari proses concatenation yang dihasilkan sebagai suara output.

Flowchart

1. *Flowchart* pada perancangan aplikasi konversi *file* teks ke audio
   Berikut adalah *flowchart* perancangan aplikasi konversi *file* teks ke audio:



Gambar 5. *Flowchart* Pada Aplikasi Teks ke Audio

Adapun tahapan pada proses diatas, yaitu:

1. Tahap Input Teks: Pengguna memasukkan teks pada kolom yang disediakan.
2. Tahap pilih Suara Pria dan Wanita : Pada tahap ini pengguna dapat memilih suara yang akan dihasilkan.
3. Proses Teks ke Suara: Teks mengalami proses konversi menjadi suara. Jika teks memiliki kurang dari 1 kata atau lebih dari 1000 kata, teks akan diulang kembali untuk proses sintesis suara.
4. Output Suara: Jika teks terdeteksi dengan benar, akan menghasilkan suara secara otomatis.
5. Refresh: Memperbarui atau mengupdate data teks untuk memulai tampilan awal atau tahap baru.
6. End: Proses selesai dalam alur kerja atau flowchart yang digambarkan.

HASIL DAN PEMBAHASAN

Pengujian

Pengujian yang dilakukan pada penetilian ini adalah *black box testing. Black box testing* untuk menguji kesesuain antara masukan yang diberikan dengan keluaran dari aplikasi [15].

Black Box Testing

Peneliti telah memberikan scenario terhadap aplikasi yang telah dibuat. Hasil dari *black box testing* ditunjukan pada Tabel 3.

Tabel 3. Hasil Black Box Testing

Skenario pengujian	**Hasil Yang Diharapkan**	**Hasil pengujian**	**Kesimpulan**
Kecepatan teks mengkonversi suara	Memerlukan 3 detik	Sesuai harapan	Valid
Mengetikkan “halo”	Menghasilkan suara	Sesuai harapan	Valid
Mengetikkan “halo”1 kata terdapat 200 huruf	Menghasilkan Suara	Sesuai, tetapi tidak dapat di simpan	Kurang Valid
Mengetikkan kata random “jaijxamjanxacjjkwnjxorjkaklkqo”	Mengahasilkan suara sesuai dengan yang di input	Sesuai harapan, dapat disimpan	Valid
Intonasi saat membaca “*Black box Testing* adalah proses mengujian perangkat lunak (software) atau aplikasi dari sudut pandang pengguna, tanpa mengetahui struktur internal atau desain struktur tersebut”.	Sesuai dengan intonasi	Sesuai harapan	Valid
Membaca tulisan Inggris “An algorithm is a set of instructions that is designed to solve a specific problem or perform a particular task. It is a well-defined procedure that takes some input and produces a corresponding output.”	Menghasilkan suara bahasa inggris	Sesuai diharapkan	Valid
Mengetikkan soal matematika “(12+8) + (-3n) = -22 20 - 3n = -22 -3n = -42”	Menghasilkan suara dengan benar	Tidak sesuai harapan, tidak dapat di simpan	Non valid
Upload *file* berupa TXT *file*	Berhasil Upload *file*	Sesuai harapan	Valid

 https://doi.org/10.56211/helloworld.v3i3.587

ImplementasiAplikasi

Tampilan Awal Antarmuka (Interface)

Gambar 6. Tampilan Awal Antarmuka *(Intreface)*

Gambar diatas merupakan tampilan yang terdapat pada Aplikasi Teks ke Audio. Pada aplikasi ini pengguna akan mengetikkan kata atau kalimat untuk menginput teks sehingga menghasilkan suara. Selain mengetikan kata atau kalimat pada aplikasi Teks ke Audio ini juga terdapat menu *choose file* yang berfungsi untuk memilih *file file* teks berupa TXT dan juga terdapat tombol suara pria dan wanita untuk memilih output yang akan di outputkan. Terdapat juga tombol *Refresh* yang berfungsi untuk mengulang kembali apabila teks ingin dihapus secara cepat.

Tampilan Teks yang Sudah Diinput

Gambar 7. Tampilan Teks Yang Sudah Di Input

Pada tampilan Gambar 7 kata atau kalimat yang sudah masuk (input). Jika aplikasi sudah benar maka system akan memperoses dengan terdengarnya keluaran (output) suara yang sesuai dengan teks yang dimasukan (input), dengan maksimal 1000 karakter. Pada proses ini, program akan mencari kata atau kalimat dengan menggunakan metode *synthesizer concatenation*. Setelah *user* melakukan aksi pengisian, *user* harus memilih suara yang akan dihasilkan. Berikut tampilan saat memilih suara pria atau wanita pada gambar 7.

Tampilan Untuk Suara Pria Atau Wanita

Gambar 8. Tampilan Memilih Suara PriaAtau Wanita

Gambar telah menginput kata atau kalimat, kemudian *user* memilih suara pria atau wanita yang akan dihasilkan.

Tampilan Saat Memilih File Teks

Gambar 9. Tampilan Saat Memilih *File* Teks

Gambar diatas ialah tampilan saat *user* memilih *file.* Teks dalam bentuk TXT *File* yang terdapat di dalam *folder* dalam PC atau Laptop. Sebagai contoh *file* 'betapa bahagianya hatiku' yang akan dipilih, lalu klik open.

Tampilan Saat Proses File Teks

Gambar 10. Tampilan Saat Proses *File* Teks

Gambar diatas ialah saat *user* telah memilih *file* teks 'betapa bahagianya hatiku', lalu memilih suara yang akan di outputkan maka secara otomatis akan adanya tampilan *button song bar.*

Tampilan Simpan Audio

Gambar 11. Tampilan Simpan Audio

Pada Gambar 11 suara yang sudah keluar (output) dapat disimpan dalam bentuk WAV.

Tampilan Saat Memilih Refresh

Gambar 12. Tampilan Saat Memilih*Refresh*

Gambar diatas ialah tampilan saat *user* memilih tombol *refresh*, maka secara otomatis tampilan akan balik seperti tampilan awal. Tombol ini berfungsi untuk mempermudah *user* jika ingin menghapus teks yang telah ada sebelumnya.

KESIMPULAN DAN SARAN

Aplikasi konversi teks ke suara yang dapat diakses melalui website mempermudah pengguna dalam mengakses layanan ini dari berbagai perangkat dengan koneksi internet. Meskipun metode synthesizer concatenation dapat mengucapkan kata-kata dalam Bahasa Indonesia dengan cukup lancar dan dapat dimengerti, pola intonasinya belum baik. Untuk menghasilkan pengucapan yang lebih natural, berintonasi, dan emosional, disarankan untuk mengembangkan metode dengan perangkaian ucapan statistik parametrik (seperti HMN) dan teknik lainnya. Menambah dukungan untuk berbagai bahasa asing seperti Inggris dan Arab, sehingga aplikasi lebih inklusif dan bermanfaat dalam konteks internasional, seperti pembelajaran bahasa, penerjemahan real-time, dan komunikasi lintas budaya. Menambahkan fitur untuk menyesuaikan intonasi, kecepatan suara, dan dialek, serta integrasi dengan layanan cloud untuk penyimpanan dan pemrosesan yang lebih cepat. Fitur tambahan seperti pengenalan konteks, emosi, dan berbagai karakter atau gaya bicara akan membuat aplikasi lebih menarik. Menggunakan Google Speech API untuk meningkatkan akurasi, mendukung berbagai bahasa dan aksen, serta memberikan kualitas suara yang lebih alami dan responsif. Ini juga memudahkan integrasi dengan berbagai platform. Menyediakan fitur perekaman langsung, pengorganisasian file yang efisien, dan kemampuan berbagi serta sinkronisasi data dengan layanan cloud. Selain itu, menambahkan alat pengeditan audio dasar dan memastikan keamanan data dengan enkripsi dan perlindungan kata sandi akan meningkatkan utilitas dan keamanan aplikasi.

DAFTAR PUSTAKA

[1] S. A. Dhiaulhaq, R. R. Ginanjar, and D. P. Lestari, "Sistem Text-to-Speech End to End Berbasis GAN untuk Bahasa Indo," no. 2, pp. 57–62, 2022.

[2] O. J. Raiyetunbi and A. Emmanuel, "An interactive cloud based user oriented, dynamic and intelligent text-to-speech module," *East African Sch. J. Eng. Comput. Sci.*, vol. 4480, no. 1, pp. 1–11, 2020, doi: 10.36349/easjecs.2020.v03i01.01.

[3] N. P. Essien, V. A. Uwah, and E. P. Ododo, "An Interactive Intelligent Web-Based Text-to-speech System for the Visually Impaired," *Asia-Africa J. Recent Sci. Res. | 76 AAJRSR*, vol. 1, p. 2021, 2021, [Online]. Available: https://journals.iapaar.com/index.php/AAJRSR

[4] M. E. Prastyo, "Aplikasi Text to Speech Berbasis Javascript," *Visualika*, vol. 7, no. 1, pp. 89–101, 2022, [Online]. Available: http://www.jurnas.stmikmj.ac.id/index.php/visualika/article/view/147/72

[5] S. Kadlag, "Konverter Teks ke Audio (TTA)," vol. 1, 2020.

[6] M. Qiraah, S. Sarif, and A. Ar, "Efektivitas Artificial Intelligence Text To Speech dalam Meningkatkan," vol. 6, no. 1, pp. 1–8, 2024, doi: 10.47435/naskhi.v6i1.2697.

[7] D. A. Firmansah, R. S. Rohman, and Y. Farlina, "Aplikasi Website Pengajuan Cuti Karyawan Rumah Sakit Islam Assyifa Sukabumi Berbasis Whatsapp Blast," *J. Teknol. dan Inf.*, vol. 10, no. 2, pp. 129–143, 2020, doi: 10.34010/jati.v10i2.2854.

[8] K. Haryodi, *Belajar Pemprograman WEB Untuk Dokumen*, Pertama. Yogyakarta: PT anak Hebat Indonesia, 2024.

[9] M. I. N. Shiddiq, "Aplikasi Untuk Mengubah Gambar Ke Teks Dan Teks Ke Suara Dengan Metode Optical Character Recognition Berbasis Android.," *UNIKOM*, no. 1, pp. 5–24, 2021.

[10] Sudirman Melangi, "Text To Speech Bahasa Indonesia Menggunakan Synthesizer Concatenation Berbasis Fonem," *J. Tek. Elektro CosPhi*, vol. 2, no. 2, pp. 31–36, 2018.

[11] N. P. Andayu, "Perancangan Text To Speech Converter Engine Dalam Pengucapan Kata Berbahasa Arab Sehari-Hari," *J. Sist. dan Teknol. Inf.*, vol. 1, no. 3, pp. 144–149, 2019.
[12] U. 2 Suska, "Implementasi dan Analisis Konversi Suara Menggunakan Algoritma Pitch Shifting dengan Time Domain Pitch Synchronous Overlap Add (TD-PSOLA)," vol. 5, no. 1, pp. 1689–1699, 2024.
[13] U. 1 Suska, "Implementasi dan Analisis Konversi Suara Menggunakan Algoritma Pitch Shifting dengan Time Domain Pitch Synchronous Overlap Add (TD-PSOLA)," pp. 18–25, 2024.
[14] Z. Rugmiaga and M. Huda, "Text Pre-Processing Pada Text To Speech Synthesis System Untuk Penutur Berbahasa Indonesia," *Ind. Electron. Semin.*, vol. 13, pp. 1–6, 2019.
[15] A. C. Praniffa, A. Syahri, F. Sandes, U. Fariha, Q. A. Giansyah, and M. L. Hamzah, "Pengujian Black Box Dan White Box Sistem Informasi Parkir Berbasis Web Black Box and White Box Testing of Web-Based Parking Information System," *J. Test. dan Implementasi Sist. Inf.*, vol. 1, no. 1, pp. 1–16, 2023.