

## Penerapan Metode *Image-to-Speech* melalui Kamera dalam Aplikasi berbasis Kecerdasan Buatan untuk Orang dengan Disleksia

Implementation of the Image-to-Speech Method via Camera in Artificial Intelligence-based Application for Individual with Dyslexia

Daniel Aprillio<sup>1</sup>, Anna Bella Atmadjaja<sup>2</sup>, Bryan<sup>3</sup>, Mychael Wijaya<sup>4</sup>, and Theresia Ratih Dewi Saputri<sup>5\*</sup>

<sup>12345</sup>Informatika, Universitas Ciputra, Surabaya, Indonesia, 60290

e-mail: <sup>1</sup>daprillio@student.ciputra.ac.id, <sup>2</sup>aatmadjaja03@student.ciputra.ac.id,

<sup>3</sup>bryan001@student.ciputra.ac.id, <sup>4</sup>msutrisna@student.ciputra.ac.id, <sup>5</sup>theresia.ratih@ciputra.ac.id

\*Corresponding author

Submitted Date: March 6<sup>th</sup>, 2024

Reviewed Date: March 10<sup>th</sup>, 2024

Revised Date: March 25<sup>th</sup>, 2024

Accepted Date: March 31<sup>th</sup>, 2024

### Abstract

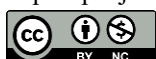
Dyslexia occurs worldwide despite the culture or language. Dyslexia affects about 9% - 12% of the population, with 2% - 4% of the population experiencing significant reading impairments. This research aims to develop an artificial intelligence-based application using the Image-to-Speech method that can convert digital text into audible sound for individuals with dyslexia without requiring their brain to process the writing. This method can assist people with dyslexia in daily life challenges such as reading traffic signs, books, or documents. Results from 10 experiments on the implementation of the proposed method indicate that individuals with dyslexia can scan the text they want to read using a camera from a smartphone or laptop. The experiments also shows that the application can convert text in image form into sound comprehensible to those with dyslexia, thus facilitating their recognition of digital writing with 90% accuracy. The application also demonstrates efficiency in terms of data processing time. The average time required for image to audio conversion is 0.22 seconds, with an average memory usage of 163.2 MiB.

Keywords: Dyslexia; computer vision; image; image-to-speech; python

### Abstrak

Disleksia terjadi di seluruh dunia terlepas dari budaya atau bahasa dan mempengaruhi sekitar 9% - 12% dari populasi, dan 2% - 4% dari populasi dapat terpengaruh secara serius keterbatasan kemampuan membaca. Penelitian ini bertujuan untuk pengembangan aplikasi berbasis kecerdasan buatan dengan metode *Image-to-Speech* yang mampu mengubah sebuah tulisan digital menjadi audio yang dapat didengar oleh orang dengan Disleksia tanpa otaknya harus memproses tulisan. Metode ini mampu membantu orang Disleksia dalam masalah kehidupan sehari-hari seperti membaca rambu lalu lintas, buku, atau dokumen. Hasil dari 10 pengujian pada penerapan metode yang diusulkan menunjukkan bahwa orang dengan dislexia dapat menscan tulisan yang ingin dibaca dengan menggunakan kamera dari handphone atau laptop. Pengujian juga menunjukkan bahwa aplikas mampu mengubah teks berupa gambar menjadi suara yang dapat dipahami oleh penderita Disleksia agar memudahkan mereka dalam mengenali tulisan secara digital dengan akurasi 90%. Aplikasi juga menunjukkan efisiensi dari segi waktu pemrosesan data. Rata-rata waktu yang dibutuhkan untuk konversi gambar menjadi audio adalah 0.22 detik dengan rata-rata penggunaan memori sebesar 163.2 MiB.

Kata Kunci: Disleksia; computer vision; gambar; image-to-speech; python



## 1. Pendahuluan

Disleksia, yang juga dikenal sebagai gangguan membaca, adalah gangguan umum yang membuat orang kesulitan membaca (Jan & Khan, 2023). Hal ini merupakan ketidakmampuan otak seseorang dalam memproses materi tertulis atau cara otak yang tidak dapat memproses materi tertulis (Martinelli & Brincat, 2022). Menurut *European Dyslexia Association* (EDA), Disleksia terjadi di seluruh dunia terlepas dari budaya atau bahasa dan mempengaruhi sekitar 9 sampai 12 persen dari populasi, dan 2 sampai 4 persen dari populasi dapat terpengaruh secara serius olehnya (Reid, 2022). Tantangan utama yang dihadapi orang dengan Disleksia yaitu kesulitan membaca yang dapat berdampak signifikan pada kinerja akademik maupun profesional mereka. Tidak hanya berdampak pada kemampuan *hard skill*, Disleksia juga dapat berdampak pada kesehatan mental seseorang (Georgiou & Parrila, 2023). Orang dengan Disleksia memiliki kecenderungan untuk berpikir negatif terhadap cara mereka belajar yang dapat berdampak negatif seperti rendahnya kemampuan akademik, *anxiety*, dan depresi (Bazen et al., 2023).

Hammill Institute on Disabilities telah melakukan sebuah simulasi dalam membaca rambu lalu lintas ketika mengemudi yang pada akhirnya, hasil menunjukkan bahwa individu dengan Disleksia menghadapi kesulitan yang signifikan dalam membaca rambu lalu lintas, salah satunya adalah tulisan digital, ketika menyetir (Tejero et al., 2019). Hal tersebut juga berlaku bukan saja kepada pengemudi tetapi semua orang yang mempunyai gangguan Disleksia. Dalam kondisi tersebut, mereka akan berada pada posisi yang kurang menguntungkan, terutama dengan rambu-rambu lalu lintas yang terpampang kata-kata yang hampir tidak pernah dibaca. Masalah seperti ini menjadi sebuah kesulitan dalam kehidupan sehari-harinya sehingga dengan adanya suatu alat yang dapat memudahkan orang dengan Disleksia untuk mengetahui konten sebuah teks akan sangat membantu. Oleh karena itu, tujuan utama dari penelitian ini adalah untuk membantu orang Disleksia dengan mudah dan cepat memahami tulisan dari berbagai objek.

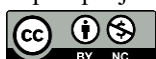
Berbagai teknologi telah dikembangkan untuk meningkatkan kemampuan membaca dan menulis bagi individu dengan Disleksia (Moncayo Arias et al., 2024). Contohnya, aplikasi dan perangkat lunak khusus telah dirancang untuk membantu meningkatkan kemampuan membaca dengan menyediakan teks dalam format yang lebih

mudah dipahami, seperti dengan penyesuaian ukuran dan jenis huruf (Surayya & Mubarak, 2021). Penelitian ini menyebutkan 90% hasil pengujian menyatakan terjadi perubahan untuk menghafal abjad. Penelitian lain menyebutkan bahwa pembelajaran membaca untuk anak dengan Disleksia meningkatkan minat dan kemambuan, khususnya untuk pemahaman huruf konsonan dobel, ketika dikemas dalam bentuk aplikasi (Moraza & Nurhastuti, 2021). Penelitian lain menyoroti bahwa aplikasi mobile dapat digunakan untuk melakukan skrining serta intervensi pada orang dengan disleksia (Politi-Georgousi & Drigas, 2020). Penelitian ini juga menyebutkan bahwa penggunaan teknologi mampu mereeducasi orang dengan disleksi untuk lebih meningkatkan kemamuan membaca dan menulis.

Teknologi kecerdasan buatan telah membantu individu dalam mengakses informasi tertulis dengan lebih mudah baik dalam ranah sosial, sains, sampai kesenian (Ko et al., 2023). Penelitian yang dilakukan oleh Husni menyatakan bahwa penting untuk ada inovasi digital yang mampu membantu anak dengan Disleksia atau orang yang memiliki kesulitan untuk membaca (Husni et al., 2023). Aplikasi lain diusulkan pada penelitian mengenai peningkatan cara belajar anak dengan disleksia (Tariq & Latif, 2016). Beberapa aplikasi tersebut masih terdapat keterbatasan khususnya dari segi ketepatan dan kecepatan. Untuk meningkatkan efektivitasnya, penelitian yang mengintegrasikan teknologi *image to text* dengan audio menjadi penting. Dengan menggabungkan kedua teknologi ini, individu dengan Disleksia dapat mendengarkan teks yang telah dikonversi dari gambar, sehingga mereka tidak hanya mengandalkan penglihatan untuk membaca, tetapi juga pendengaran.

Teknologi *image-to-speech* merupakan teknologi kecerdasan buatan yang mampu mengkonversi sebuah gambar menjadi baris kata dan kemudian dinarasikan dalam bentuk suara (Sheffer & Adi, 2023). Penggunaan *image-to-speech* sendiri sudah banyak digunakan. Sayangnya, belum ada aplikasi spesifik yang dapat digunakan untuk orang dengan Disleksia. Program penerapan *text-to-speech* masih terbatas dalam hal akurasi, fleksibilitas, dan kemudahan penggunaan.

Berdasarkan uraian tersebut, peneliti mengusulkan penerapan metode *Image-to-Speech* yang mampu mengubah sebuah tulisan digital menjadi audio yang dapat didengar oleh penderita Disleksia tanpa otaknya harus memproses tulisan. Penelitian ini bertujuan untuk membantu orang

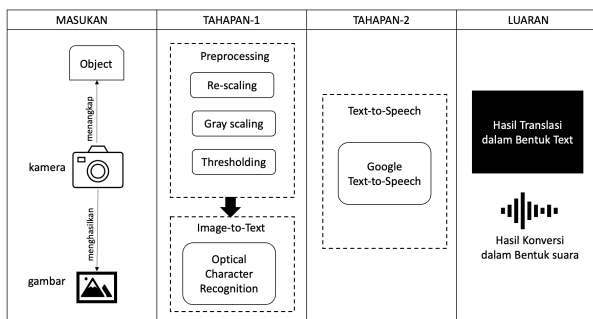


dengan disleksia dengan menggunakan metode kecerdasan buatan. Dengan kebaruan penelitian ini, aplikasi yang dikembangkan diharapkan mampu membantu orang dengan Disleksia dalam masalah kehidupan sehari-hari seperti membaca rambu lalu lintas, buku, atau dokumen karena otak mereka yang tidak dapat memproses tulisan. Orang dengan Dislexia dapat men-*scan* tulisan yang ingin dibaca dengan menggunakan kamera dari handphone atau laptop. Sistem akan mengubah teks berupa gambar menjadi suara yang dapat dipahami oleh penderita Disleksia agar memudahkan mereka dalam mengenali tulisan secara digital dengan akurasi yang tinggi. Hal ini dapat berdampak pada peningkatan kemudahan pada orang dengan Disleksia untuk mengakses dan mendapat pemahaman yang komprehensif dari informasi yang sifatnya tertulis.

## 2. Metode Penelitian

Penelitian ini menerapkan metode *image-to-speech* yang dikembangkan dengan menggunakan bahasa pemrograman Python. Bahasa pemrograman ini dipilih karena sudah memiliki berbagai *library* yang telah dikembangkan terkhususnya untuk pengembangan sistem berbasis kecerdasan buatan. Dalam menerapkan metode *image-to-speech* tersebut, terdapat dua tahapan utama yang harus dilakukan untuk memproses tulisan dalam bentuk gambar menjadi suara seperti yang terlihat pada Gambar 1.

Hasil luaran penelitian selanjutnya diuji dengan menggunakan *Black Box testing* (Sasmito & Nishom, 2020). Pengujian ini bertujuan untuk melihat fungsionalitas sistem secara keseluruhan. Pengujian *black box* dipilih karena sistem masih digunakan dalam lingkungan laboratorium dan belum diujikan secara umum.



Gambar 1. Metode Penelitian

### a. Tahapan-1: Pemrosesan Gambar

Tahap pertama adalah memproses gambar yang diterima oleh sistem menjadi bentuk tulisan

digital yang dapat dibaca oleh sistem. Adapun beberapa *library* di Python yang bisa digunakan untuk membantu memecahkan masalah pada tahap pertama ini. *Library* pertama adalah OpenCV. *Library* tersebut dapat digunakan untuk membantu sistem mendapat informasi berupa tulisan pada gambar yang didapatkan dari kamera (Bradski et al., 2000). OpenCV dipilih karena kinerja baiknya dalam mengolah data berupa gambar. *Library* ini banyak digunakan pada penerapan computer vision di berbagai bidang (Kumar et al., 2023).

Setiap gambar yang diterima perlu melalui *preprocessing* terlebih dahulu. *Preprocessing* bertujuan untuk menyesuaikan gambar agar lebih mudah dibaca oleh sistem pada proses pengambilan informasi (Suresh & Niranjnamurthy, 2021). Tahap pertama pada *preprocessing* adalah melakukan *resizing*. Setiap gambar yang dimasukkan oleh pengguna akan sangat bervariasi kualitasnya. Dengan melakukan *resizing* menggunakan fungsi “*resize()*” dari OpenCV, gambar akan diperbesar agar memiliki DPI (Dots per Inch) di atas 300 DPI. Ketika sistem sudah dapat membaca gambar, selanjutnya sistem perlu untuk mengubah gambar berwarna menjadi citra biner. Hal ini perlu dilakukan untuk menyederhanakan algoritma yang akan digunakan untuk mengambil tulisan dari gambar. Untuk mencapai hal tersebut, salah satu fungsi dari *library* OpenCV yaitu “*cvtColor()*” dapat digunakan untuk mengubah *color space* dari suatu gambar ke *color space* yang lainnya. Dalam sistem ini, maka gambar berwarna akan diubah menjadi *grayscale* (Kanan & Cottrell, 2012). Implementasi tahapan pertama ini dilakukan sesuai dengan *pseudocode* yang tertera pada Gambar 2.

Algorithm 1 Image Processing

```

1: function CAPTUREIMAGE
2:   camera ← OPENCAMERA
3:   frame ← CAPTUREFRAME(camera)
4:   imagePath ← “captured_image.jpg”
5:   SAVEIMAGE(frame, imagePath)
6:   RELEASECAMERA(camera)
7:   return imagePath
8: end function
9: function PROCESSIMAGE(imagePath)
10:  image ← READIMAGE(imagePath)
11:  resizedImage ← RESIZE(image, 300 dpi)
12:  grayscaleImage ← CVTCOLOR(resizedImage, GRAYSCALE)
13:  THRESHOLD(grayscaleImage, THRESH_BINARY, THRESH_OTSU)
14:  return grayscaleImage
15: end function
16: function PERFORMOCR(processedImage)
17:  text ← EXTRACTTEXTWITHOCR(processedImage)
18:  return text
19: end function
    
```

Gambar 2. Pseudocode untuk Pemrosesan Gambar

Setiap gambar yang sudah menjadi *grayscale* kemudian akan melalui proses

*thresholding*. Pixel yang memiliki nilai di atas *threshold* akan diubah menjadi pixel putih atau dianggap sebagai bukan tulisan (Bhahri & others, 2018). Sedangkan, *pixel* yang memiliki nilai di bawah *threshold* akan diubah menjadi pixel hitam atau dianggap sebagai tulisan. Proses *thresholding* ini juga menggunakan fungsi “*threshold()*” dengan menggunakan teknik “*THRESH\_BINARY*” sesuai dengan aturan di atas. Namun, tidak semua gambar memiliki tingkat pencahayaan yang sama. Ketika hal tersebut terjadi, teknik di atas akan menjadi kurang akurat. Oleh karena itu, teknik “*THRESH\_OTSU*” juga digunakan untuk menentukan nilai *threshold* secara otomatis. Dengan seluruh proses tersebut, gambar yang akan diproses selanjutnya merupakan citra biner.

Setelah melakukan *thresholding*, sistem perlu mengetahui area pada citra biner yang memiliki tulisan dengan fungsi “*getStructuringElement()*” yang dapat menentukan luas area yang perlu diperiksa pada proses selanjutnya dengan menghubungkan bagian-bagian yang terputus yang dalam hal ini menggunakan bentuk kotak dengan menggunakan teknik “*MORPH\_RECT*”. Setelah area bagian terluar tulisan diketahui dalam bentuk titik-titik, titik-titik tersebut dapat diperluas dengan fungsi “*dilate()*”. Hal ini berfungsi untuk menebalkan area bagian terluar tulisan yang telah didapat sebelumnya sehingga tulisan terlihat lebih tebal dan jelas.

Selanjutnya, dilakukan pengenalan karakter pada gambar dengan metode *Optical Character Recognition* (OCR). Metode ini merupakan metode yang banyak digunakan untuk mengkonversi dokumen digital dalam bentuk yang dapat dibaca oleh mesin/komputer (Nguyen et al., 2021). Pengambilan informasi berupa tulisan menjadi bentuk digital dilakukan dengan menggunakan *library Tesseract*. *Library* ini dikembangkan oleh Google sebagai OCR atau *Optical Character Recognition* untuk mendeteksi tulisan pada perangkat. Namun, berkaitan sistem yang akan dibuat menggunakan Python, maka *library* yang akan digunakan adalah *PyTesseract* yang dibuat berdasarkan *Tesseract*. *PyTesseract* bekerja ketika citra yang diberikan memiliki tulisan berwarna hitam dengan bagian lain dari citra berwarna putih (Dome & Sathe, 2021). Untuk mencapai hal tersebut, sistem akan menukar seluruh *pixel* yang ada pada citra. Selanjutnya, penggunaan salah satu fungsi dari *PyTesseract* yaitu “*image\_to\_string()*”, hasil yang sudah didapat sebelumnya akan diubah menjadi tulisan digital.

## b. Tahapan-2: Konversi Teks ke Suara

Tahap kedua adalah mengelola informasi tulisan digital yang telah diterima menjadi audio. Tetapi sebelum itu, tulisan digital yang terbaca oleh *PyTesseract* dapat berasal dari bahasa apapun sesuai dengan bahasa yang dapat dideteksi oleh *Tesseract*. Untuk dapat mengubah informasi tulisan digital ke dalam bentuk audio, maka sistem perlu mengetahui bahasa dari tulisan tersebut sehingga sistem dapat mengeluarkan audio yang sesuai. Hal ini dapat dicapai dengan menggunakan *library googletrans* (Cire, 2021). *Library* ini akan mendeteksi bahasa yang digunakan dalam tulisan digital dengan menggunakan fungsi “*detect()*”.

Setelah bahasa dideteksi, *library gTTS* (Google Text-to-Speech) akan mempermudah pengelolaan tulisan digital menjadi audio (Iancu, 2019). Dengan fungsi yang disediakan oleh *gTTS*, sistem hanya perlu memberikan informasi berupa tulisan digital yang ingin diubah ke dalam bentuk audio, bahasa tulisan, dan *boolean* yang menandakan apabila tulisan harus dibacakan secara lambat. Audio yang diterima akan disimpan terlebih dahulu ke dalam format audio yaitu *mp3* sebelum diputar oleh sistem melalui *output* audio yang tersedia di perangkat pengguna. Proses ini diimplementasikan sesuai dengan *pseudocode* yang tertera pada Gambar 3. Program dijalankan dengan memanggil semua fungsi di masing-masing algoritma seperti yang tertera pada *pseudocode* pada gambar 4. Proses akan diawali dengan mengambil gambar dengan kamera, memproses gambar, mengenali teks pada gambar, dan melafalkan teks menjadi audio.

### Algorithm 2 Text to Speech Conversion

```
1: function TEXTTOSPEECH(text)
2:   speech ← CONVERTTEXTTOSPEECH(text)
3:   audioPath ← "speech.mp3"
4:   SAVEAUDIO(speech, audioPath)
5:   PLAYAUDIO(audioPath)
6: end function
```

Gambar 3. *Pseudocode* untuk mengubah Teks menjadi Audio

### Algorithm 3 Main Process

```
1: function MAIN
2:   imagePath ← CAPTUREIMAGE
3:   processedImage ← PROCESSIMAGE(imagePath)
4:   text ← PERFORMOCR(processedImage)
5:   TEXTTOSPEECH(text)
6: end function
7: MAIN
```

Gambar 4. *Pseudocode* untuk Proses Utama

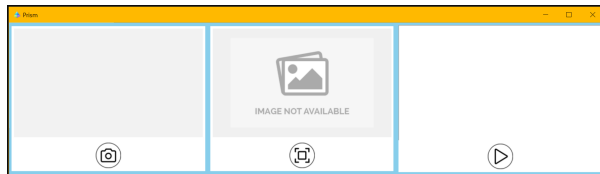
## 3. Hasil dan Pembahasan

Hasil implementasi dari metode yang diusulkan dibuat dalam bentuk aplikasi *desktop*.

Selanjutnya, aplikasi diuji dengan 10 skenario pengujian dengan jenis inputan yang berbeda-beda.

#### a. Deskripsi Program

Implementasi *user interface* (UI) pada aplikasi untuk penelitian ini dilakukan dengan menggunakan *Tkinter* (Tirtana et al., 2021). Ketika program dijalankan, antarmuka pengguna grafis akan ditampilkan seperti pada Gambar 5 dibawah ini.



Gambar 5. Tampilan Utama Program

Pada *frame* pertama di bagian kiri, camera berperan sebagai sensor untuk menangkap informasi yang ingin diproses yaitu, tulisan digital atau environment. Ketika pengguna klik tombol kamera, maka gambar akan tersimpan pada folder aplikasi. Pada *frame* kedua yaitu pada bagian tengah, hasil gambar yang ditangkap akan ditunjukkan agar dapat dicek kualitas foto oleh para pengguna, jika kualitas foto dianggap kurang memuaskan maka pengguna dapat mengambil gambar ulang untuk diproses. Kemudian terdapat tombol scan yang dapat digunakan pengguna untuk menjalankan program yang mengubah gambar menjadi *text*. Kamera dapat mengambil gambar dari berbagai macam jenis inputan seperti buku, menu makanan, rambu lalu lintas, dan lain-lain. Hal ini diharapkan dapat meningkatkan efektifitas dan efisiensi pada sistem yang dibangun. Sayangnya, masih ada batasan dalam aplikasi ini yaitu gambar hanya dapat diambil dari benda yang sifatnya fisik. Hal ini menyebabkan pengguna tidak dapat mengambar gambar dalam halaman website, kecuali jika halaman website tersebut dibuka menggunakan gawai yang lain.

Program tersebut berupa perubahan gambar berwarna menjadi citra biner. Hal ini perlu dilakukan untuk menyederhanakan algoritma yang akan digunakan untuk mengambil tulisan dari gambar yaitu *thresholding*. Setelah itu, sistem perlu mengetahui area pada gambar yang memiliki tulisan. dengan menentukan luas area yang perlu diperiksa dengan menghubungkan bagian-bagian yang terputus, hasil yang didapat akan disimpan dalam bentuk koordinat Numpy berupa array dari sumbu x dan y. Pengambilan tulisan menjadi bentuk digital dari kotak-kotak yang berisikan

tulisan dilakukan menggunakan library Tesseract. Kemudian hasil outputnya akan berupa file dalam bentuk *.txt*.

Pada *frame* terakhir di bagian kanan, hasil file *.txt* akan ditampilkan, tentunya bagi pengguna dengan Disleksia tidak dapat membaca tulisan tersebut tetapi fitur ini ditambahkan sebagai konfirmasi bahwa image yang di scan telah berhasil diubah menjadi file *.txt*. Kemudian terdapat tombol play audio yang berfungsi untuk mengubah *text* menjadi audio atau *actuators* yang akan didengar oleh para pengguna. Dengan fungsi yang disediakan oleh *gTTS*, sistem hanya perlu memberikan informasi berupa tulisan digital yang ingin diubah ke dalam bentuk audio, bahasa tulisan, dan boolean yang menandakan apabila tulisan harus dibacakan secara lambat. Hal tersebut kemudian dapat dinilai akurasi performanya.

#### b. Hasil Pengujian

Proses utama aplikasi terbagi dalam tiga bagian utama yaitu menangkap gambar dari kamera, konversi gambar ke teks, dan konversi teks ke audio. Performa serta kontribusi dari penelitian ini dievaluasi melalui proses penangkapan gambar diuji berdasarkan penggunaan memori dan waktu yang dibutuhkan.

Proses pengujian *blackbox* dilakukan dengan menggunakan beberapa gambar yang berbeda. Dengan mengevaluasi gambar-gambar yang berbeda konteks ini, penelitian ini mampu menunjukkan bahwa metode yang diusulkan mampu digunakan dalam situasi yang berbeda-beda.

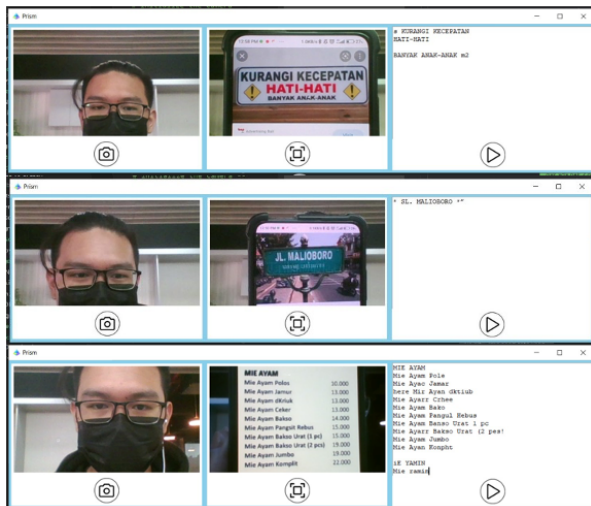
Pada satu kali uji coba proses penangkapan gambar, didapatkan bahwa rata-rata penggunaan memorinya untuk satu kali proses adalah 158.7678 MiB. Waktu yang dibutuhkan untuk satu kali proses adalah 0.0189 detik. Setelah melakukan uji coba proses penangkapan gambar sebanyak 10 kali, didapatkan bahwa rata-rata penggunaan memorinya untuk 10 kali proses adalah 158.5760 MiB. Rata-rata waktu yang dibutuhkan untuk 10 kali proses adalah 0.0145 detik.

Performa proses konversi gambar ke teks diuji berdasarkan penggunaan memori dan waktu yang dibutuhkan. Pada satu kali uji coba konversi gambar ke teks, didapatkan bahwa rata-rata penggunaan memorinya untuk satu kali proses adalah 162.0898 MiB. Waktu yang dibutuhkan untuk satu kali proses adalah 0.1934 detik. Setelah melakukan uji coba konversi gambar ke teks sebanyak 10 kali, didapatkan bahwa rata-rata penggunaan memorinya untuk 10 kali proses

adalah 162.4080 MiB. Rata-rata waktu yang dibutuhkan untuk 10 kali proses adalah 0.1858 detik.

Performa proses konversi teks ke audio diuji berdasarkan penggunaan memori dan waktu yang dibutuhkan. Pada satu kali uji coba konversi teks ke audio, didapatkan bahwa rata-rata penggunaan memorinya untuk satu kali proses adalah 159.7969 MiB. Waktu yang dibutuhkan untuk satu kali proses adalah 0.5644 detik. Setelah melakukan uji coba konversi teks ke audio sebanyak 10 kali, didapatkan bahwa rata-rata penggunaan memorinya untuk 10 kali proses adalah 163.2012 MiB. Rata-rata waktu yang dibutuhkan untuk 10 kali proses adalah 0.2210 detik.

Selanjutnya, pengujian dilakukan pada 10 jenis teks yang berbeda seperti rambu lalu lintas, menu makan, dan papan nama jalan. Beberapa contoh pengujian dapat dilihat pada Gambar 6.

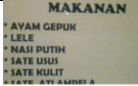



Gambar 6. Pengujian Aplikasi

Tabel 1 menunjukkan hasil pengujian dengan 10 skenario yang berbeda. Berdasarkan 10 pengujian yang telah dilakukan, dapat dilihat hasil evaluasinya bahwa rata-rata akurasi Image to Text adalah 91.1% dan rata-rata akurasi Text to Speech adalah 100%. Maka berdasarkan target yang ingin dicapai ketika memulai proyek ini yaitu 90% akurasi, dapat dilihat bahwa program telah berjalan dengan baik dan memenuhi ekspektasi.

Tabel 1 Skenario dan Hasil Pengujian

Gambar	Hasil	Akurasi Image-to-Text	Akurasi Text-to-Speech
	HATI-HATI !!! KURANGI KECEPATAN	100%	100%
	BB KURANG! KECEPATAN SEKARANG Rawan Kecelakaan	97%	100%
	KURANGI KECEPATAN HATI-HATI BANYAK ANAK-ANAK	100%	100%
	PELAN BANYAK Anak REQ HAT HAT DILARANG NGEBUT ; BANYAK ANAK-BND SL.	75%	100%
	MALIOBORO	91%	100%
	MIE AYAM Mie Ayam Pole Mie Ayam Jamar here Mir Ayan dktiub Mie Ayarr Crhee Mie Ayam Bako Mie Ayam Pangul Hebus Mie Ayam Bango Urat 1 pc Mie Ayarr Bakso Urat (2 pes! Mie Ayam Jumbo Mie Ayan Komphit iE YAMIN Mie ramin	86%	100%
	MH DOANE Mie AvAH 1008 MIE AYAM JuMBa . 10 00 MIE AYAM Pans RP 7190 HIE AVAM Bax P4000 MIL AVAM KOMPLIT PANGIIT BAKO Roa ASO PORS! BANCS! ? GO KULIT GORENG (HIE	75%	100%

Gambar	Hasil	Akurasi Image-to-Text	Akurasi Text-to-Speech
	DAFTAR MENU MAKANAN AVAM GEPUK LELE NASI PUTIH SATE USUS	90%	100%
	Pd. Kelapa Pd. Kopi Jatibening JL. WONOCOLO >ABRIK KULIT KODE POS : 60237	100%	100%
	JL. WONOCOLO >ABRIK KULIT KODE POS : 60237	97%	100%

#### 4. Kesimpulan

Disleksia yang merupakan gangguan pada otak yang memproses bahasa sehingga mereka melihat tulisan seperti acak-acakan maupun kalimat yang tidak lengkap. Penelitian mampu menjawab kesulitan orang dengan Disleksia mereka agar otak yang memproses bahasa tidak perlu lagi digunakan. Hal ini dikarenakan Disleksia digolongkan menjadi penyakit yang tidak dapat disembuhkan.

Pengujian yang dilakukan menunjukkan bahwa dengan penelitian ini, pengguna dengan Disleksia dapat mengurangi potensi salah mengucapkan kata atau salah pengertian ketika membaca sebuah tulisan digital. Dengan rata-rata akurasi sebesar 90%, penelitian ini mampu mengenali kata-kata pada gambar hasil foto dan menampilkannya dalam bentuk audio atau suara.

Hal ini dapat meningkatkan keinginan penderita Disleksia dalam membaca yang dapat membuat dirinya mendapatkan pengetahuan dan wawasan yang lebih luas. Sehingga dinantikan dimana saatnya para penderita Disleksia yang tidak suka membaca akan senang membaca guna masa depan mereka dan dunia.

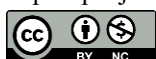
#### 5. Saran Pengembangan

Walaupun penelitian ini menunjukkan bahwa teknologi, secara khusus *computer vision*, dapat membantu pengidap Disleksia, namun hasil penerapan teknologi belum efisien. Hal ini dikarekan aplikasi hanya dapat digunakan pada laptop atau komputer. Pengembangan lebih lanjut dapat dilakukan dengan menambahkan platform mobile serta melakukan efisiensi algoritma. Selain itu terdapat keterbatasan pada jenis inputan. Saat ini penelitian ini hanya mampu menangani inputan berupa barang fisik seperti buku, menu, atau gawai

yang lain. Oleh karena itu, pengembangan kedepan termasuk kemampuan sistem untuk mengambil screenshot dari suatu halaman website atau aplikasi yang lain.

#### Referensi

- Bazen, L., de Bree, E. H., van den Boer, M., & de Jong, P. F. (2023). Perceived negative consequences of dyslexia: the influence of person and environmental factors. *Annals of Dyslexia*, 73(2), 214–234.
- Bahri, S., & others. (2018). Transformasi Citra Biner Menggunakan Metode Thresholding Dan Otsu Thresholding. *E-JURNAL JUSITI: Jurnal Sistem Informasi Dan Teknologi Informasi*, 7(2), 196–203.
- Bradski, G., Kaehler, A., & others. (2000). OpenCV. *Dr. Dobb's Journal of Software Tools*, 3(2).
- Cire, R. (2021). *A Multilingual Scrabble Game using the Googletrans Library in Python*.
- Dome, S., & Sathe, A. P. (2021). Optical character recognition using tesseract and classification. *2021 International Conference on Emerging Smart Computing and Informatics (ESCI)*, 153–158.
- Georgiou, G., & Parrila, R. (2023). Dyslexia and mental health problems. *Encyclopedia of Mental Health*, 3.
- Husni, H., Nasri, N. I. S. M., & Saip, M. A. (2023). Eye-Tracking Usability Data of BacaDisleksia for an Informed Dyslexia-Friendly Design Decision. *International Conference on Computing and Informatics*, 69–80.
- Iancu, B. (2019). Evaluating Google speech-to-text API's performance for Romanian e-learning resources. *Informatika Economica*, 23(1), 17–25.
- Jan, T. G., & Khan, S. M. (2023). A systematic review of research dimensions towards dyslexia screening using machine learning. *Journal of The Institution of Engineers (India): Series B*, 104(2), 511–522.
- Kanan, C., & Cottrell, G. W. (2012). Color-to-grayscale: does the method matter in image recognition? *PloS One*, 7(1), e29740.
- Ko, H.-K., Park, G., Jeon, H., Jo, J., Kim, J., & Seo, J. (2023). Large-scale text-to-image generation models for visual artists' creative works. *Proceedings of the 28th International Conference on Intelligent User Interfaces*, 919–933.
- Kumar, A., Samal, S., Saluja, M. S., & Tiwari, A. (2023). Automated Attendance System Based



- on Face Recognition Using Opencv. *2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)*, 1, 2256–2259.
- Martinelli, V., & Brincat, B. (2022). The similarity of phonological skills underpinning reading ability in shallow and deep orthographies: a bilingual perspective. *International Journal of Bilingual Education and Bilingualism*, 25(6), 2095–2108.
- Moncayo Arias, M. A., Bastidas Vera, E. A., Cabezas Macías, P. M., Ledesma Espin, C. del R., & Bayas Guevara, B. I. (2024). *Innovative and Inclusive Digital Applications to Enhance Literacy in Students with Dyslexia*.
- Moraza, A., & Nurhastuti, N. (2021). Mengurangi Kesalahan Membaca Permulaan Pada Anak Disleksia (X) Melalui Media Pembelajaran Berbasis Aplikasi Game Secil. *Jurnal Penelitian Pendidikan Khusus*, 9(1), 35–43.
- Nguyen, T. T. H., Jatowt, A., Coustaty, M., & Doucet, A. (2021). Survey of post-OCR processing approaches. *ACM Computing Surveys (CSUR)*, 54(6), 1–37.
- Politi-Georgousi, S., & Drigas, A. (2020). *Mobile Applications, an Emerging Powerful Tool for Dyslexia Screening and Intervention: A Systematic Literature Review*.
- Reid, G. (2022). Dyslexia around the Globe: Perspectives on Practice. In *The Routledge International Handbook of Dyslexia in Education* (pp. 386–396). Routledge.
- Sasmito, G. W., & Nishom, M. (2020). Testing the Population Administration Website Application Using the Black Box Testing Boundary Value Analysis Method. *2020 IEEE Conference on Open Systems (ICOS)*, 48–52.
- Sheffer, R., & Adi, Y. (2023). I hear your true colors: Image guided audio generation. *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1–5.
- Surayya, S., & Mubarak, H. (2021). Pengaruh Aplikasi Marbel Membaca Terhadap Kemampuan Membaca Anak Disleksia. *Jurnal Ibriz: Jurnal Kependidikan Dasar Islam Berbasis Sains*, 6(2), 193–204.
- Suresh, H. S., & Niranjanamurthy, M. (2021). Image Processing Using OpenCV Technique for Real World Data. *Intelligent Computing Paradigm and Cutting-Edge Technologies: Proceedings of the Second International Conference on Innovative Computing and Cutting-Edge Technologies (ICICCT 2020)*, 285–296.
- Tariq, R., & Latif, S. (2016). A mobile application to improve learning performance of dyslexic children with writing difficulties. *Journal of Educational Technology & Society*, 19(4), 151–166.
- Tejero, P., Insa, B., & Roca, J. (2019). Difficulties of drivers with dyslexia when reading traffic signs: Analysis of reading, eye gazes, and driving performance. *Journal of Learning Disabilities*, 52(1), 84–95.
- Tirtana, E., Gunadi, K., & Sugiarto, I. (2021). Penerapan Metode YOLO dan Tesseract-OCR untuk Pendataan Plat Nomor Kendaraan Bermotor Umum di Indonesia Menggunakan Raspberry Pi. *Jurnal Infra*, 9(2), 241–247.