

Identifikasi Teks dari Citra Menggunakan Optical Character Recognition

Akhmad Fadjeri¹

¹Teknik Informatika, Universitas Ma'arif Nahdlatul Ulama Kebumen

¹ akhmadfadjeri@umnu.ac.id

ABSTRACT

Writing is a human activity to gather information. Now the medium of writing is not only paper but also digital writing on computers. Persons who work as writers often come across documents in which there is quite a lot of text. Input text into the computer; of course, it takes a long time if typed manually one by one. Computer science, there is a field of science called image processing; in this case, it can be used as a method to make it easier in these conditions. Generally, to input text into the computer, the author handles it manually with the keyboard, but with image processing, we can process an image into a text that can be directly processed by the computer. These problems is a program was created that can convert an image into text using the Python programming language. Convert the image into text using the tesseract module. Tested using 10 sample images taken from the internet and camera photos. The results obtained show that programming using image processing with the Python programming language is able to convert an image into text with an accuracy level of 76%, the Tesseract module can detect text in images.

Kata kunci : OCR, Identifikasi Teks , Pengolahan citra digital.

I. PENDAHULUAN

Menulis adalah kegiatan manusia untuk memberikan informasi kedalam sebuah catatan. Manusia modern menggunakan huruf atau lazim dikenal dengan alfabet untuk merangkai sebuah tulisan (Y. et al., 2021). Sebelum adanya huruf manusia di jaman dahulu mengekspresikan keinginannya kedalam sebuah lukisan yang di buat di dinding goa dengan arang atau goresan dari tulang maupun batu. Seiring kemajuannya manusia mulai mengenal simbol simbol, dari simbol inilah manusia merangkainya menjadi sebuah huruf. Untuk dapat menulis manusia membutuhkan sebuah media, dimana di jaman dahulu manusia menggunakan bahan yang disediakan oleh alam seperti batu, daun, kulit, dan sebagainya (Kembarawan, n.d.). Dikarenakan media yang tersedia di alam mudah lekang atau gampang rusak manusia kemudian membuat sebuah media yang lebih terbarukan dan lebih awet yaitu kertas sebagai media tulis konvensional dan komputer sebagai media tulis digital.

Penggunaan komputer semakin pesat di bidang kecerdasan buatan (BURHAEIN et al., 2023), salah satunya teknologi media tulis. Pada awalnya untuk memasukan tulisan ke perangkat CPU dengan cara diketik, sekarang telah banyak cara memasukan tulisan kedalam piranti CPU salah satunya voice to text atau dengan menginput suara dari user lalu mengubahnya menjadi output teks, Pengguna hanya perlu bersuara mengucapkan kata yang mereka inginkan. Sebagai contoh penerapan metode ini dijalankan pada google translate.

Terdapat berbagai model memasukan tulisan suara, model terbarunya pengguna dapat menambahkan tulisan yang bersumber dari *image* kedalam CPU. Permasalahan yang dihadapi saat ini dalam penginputan teks manual dari gambar yang diketik oleh manusia (Rapar et al., 2023). Model ini mulai di kembangkan dan banyak digunakan karena dengan metode ini dapat mempercepat pekerjaan dalam menulis sehingga lebih efisien dari pada harus mengetik manual memerlukan waktu yang lebih lama dan

terkadang terdapat kesalahan dalam pengetikan (Sugiarta et al., 2021). Pada penelitian ini akan membuat sebuah program dimana tujuannya adalah mengenali atau mengidentifikasi teks yang bersumber dari citra atau gambar yang bersumber dari internet maupun hasil foto, kemudian dengan metode image processing program akan mengekstrak gambar atau citra tersebut dan menginputnya menjadi sebuah teks ke dalam komputer.

II. KAJIAN PUSTAKA

Penelitian yang peneliti lakukan di dukung oleh penelitian terdahulu yang relevan. Penelitian yang berjudul Ekstraksi Informasi/Data e-KTP Menggunakan Optical Character Recognition CNN oleh (Sugiarta et al., 2021), penelitian selanjutnya yang berjudul Analisis membandingkan OCR di Google dan di Microsoft dengan objek ktp-el (Valentino & Susetyo, 2023), layanan pendataan di Pertamina lahendong telah melakukan penelitian menggunakan OCR (Rapar et al., 2023). Penelitian terkait plat motor kendaraan roda dua untuk keamanan parkir di Universitas Indo Global (Fernandes et al., 2023). Urgensi dari penelitian yang peneliti lakukan dari penelitian terdahulu yang berfokuskan terhadap kata bukan terhadap teks.

2.1 Citra

Citra merupakan gambaran yang di keluarkan dari hasil analisis baik manual dengan mata manusia atau dengan alat rekam optik kamera dan sejenisnya, (Rizky pratama, 2021) Citra yang di dihasilkan dari pengamatan dapat berupa foto atau gambar diam maupun video atau gambar yang bergerak. (Arini et al., 2020). Objek pada sebuah citra memiliki karakternya sendiri sendiri, itulah yang membedakan citra satu dengan lainnya. Lalu bagaimana sebagai pengamat dalam hal ini manusia mampu menginterpretasikan sebuah citra, munculah metode untuk mengidentifikasi objek pada sebuah citra atau di kenal sebagai pengolahan citra. Umumnya untuk piranti yang di pakai dalam pengolahan citra adalah computer (Mohan & Poobal, 2018). Gambar atau citra yang di olah di CPU inilah yang biasa dikatakan citra digital. Citra dimasukan dalam CPU menjadi bentuk numeric sebagai intensitas besar kecilnya sebuah citra pada masing-masing piksel, dimana setiap piksel digambarkan sebagai satu kotak kecil yang memiliki koordinat posisi yaitu x menyatakan menunjukkan posisi kolom dan y menunjukkan posisi baris (Ripai & Saputri, 2023)

2.2 Pengolahan Citra Digital

Image Processing atau dalam bahasa indonesia di sebut sebagai diolahnya sebuah gambar merupakan teknik untuk memperbaiki kualiti dari gambar dengan meningkatkan (kontras, menguba warna, serta merestorasi citra), dengan mentransformasinya menjadi (sebuah skala, geometri, rotasi transformasi , translasi), atau serta menentukan ciri sebuah citra yang optimat dengan tujuan menganalisa, *save* data telah di reduksi dan pengkompresian, menantransmisi data & waktu pemrosesan (Munantri et al., 2020). Dengan kata lain proses dalam pengolahan citra adalah berupa menginput citra dan menghasilkan output juga berupa citra namun yang memiliki kualitas yang lebih baik. (Antara et al., 2023). Tidak hanya sebatas itu saja pengolahan citra juga memiliki manfaat diantaranya mengurangi ukuran file dari sebuah citra, mengubah citra kepada kondisi asal, dan ekstrasi citra untuk tujuan analisis (Fadjeri et al., n.d.). Untuk mengolah sebuah citra , komputer pada nantinya akan mengubah citra asli kedalam citra yang lain agar dimengerti oleh sistem dengan cara mengelompokkannya kedalam beberapa penggolongan citra. Berdasarkan warnanya citra di golongkan menjadi citra RGB, citra grayscale, dan citra biner.(Fadjeri, Rahmawati, et al., 2022)

2.3 Citra Warna

Citra warna dengan kata lain *color image* yang berisikan nilai RGB adalah citra yang terbentuk dari 3 warna yaitu merah, hijau dan biru. Jika warna tersebut di kombinasikan bersama akan menghasilkan beragam warna. Di dalam komputer citra RGB di repretasikan ke dalam sebuah piksel, semakin tinggi jumlah piksel yang tersedia maka citra atau gambar yang di tampilkan semakin tajam dalam artinya perpaduan warna semaikin jelas (Dan et al., 2020).

Teori Visi tri stimulus menyatakan manusia melihat warna berpijakan perbandingan cahaya yang dibandingkan dengan yang datang setelahnya dan diterima sensor sensitif cahaya di mata. Sensor tersebut peka dengan cahaya yang panjang gelombangnya 630 nm (*Red*), 530 nm (*Green*) serta 450 nm (*Blue*) (Fadjeri, Saputra, et al., 2022).

2.4 Citra Grayscale

Citra grayscale merupakan citra yang tersusun oleh satu warna yaitu abu-abu. Di dalam koputer untuk merepretasikan citra ini ke dalam piksel di dasarkan pada derajat keabuannya. Tujuan utama mengubah citra warna kedalam citra grayscale adalah memudahkan komputer untuk mengidentifikasi dengan mengurangi informasi yang tidak di perlukan di karenakan grayscale hanya berisikan informasi pencahayaan (Nabusa, 2019) Pencahayaan maksimal di repretasikan kedalam warna putih sementara pencahayaan minimum adalah hitam, serta mengidentifikasikan citra diantaranya adalah yaitu abu abu.

2.5 Optical Character Recognition

Optical Character Recognition adalah aplikasi *computer* yang bisa menerjemahkan, mengolah atau mengekstrak tulisan dengan membaca karakters pada gambar baik dari alat *print* maupun tulis manual manusia kedalam bentuk *text* yang dapat di baca atau sunting oleh komputer. (Effendi et al., 2019) Tujuannya adalah menerjemahkan bacaan yang dapat di baca oleh manusia kemudian mengubahnya menjadi sebuah kode yang dapat dimengerti oleh computer (Kurniawan et al., 2023) Untuk menggunakan metode ini citra yang di olah adalah citra yang di dalamnya terdapat karakter atau teks yang bersal dari alat cetak (printer atau mesin ketik), tulisan tangan, atau dengan cara memindai dengan kamera , atau dengan alat scan.

3. METODE PENELITIAN

3.1 Jenis Penelitian

Dalam melaksanakan penelitian ini metode yang di pakai adalah metode eksperimental. Metode ini digunakan untuk mengetahui sebab akibat. Tahapan tahapan dari metode ini dimulai dari memilih topik penelitian, merumuskan masalah, merumuskan hipotesis, menentukan variabel, merancang program, melaksanakan penelitian, menganalisis hasil penelitian, dan langkah terakhir adalah membuat kesimpulan.

3.2 Objek Penelitian

Objek penelitian menjadi perhatian utama dimana menjadi rujukan dalam mendapatkan sebuah jawaban maupun solusi dari permasalahan yang dihadapi. Objek pada penelitian ini menggunakan dua jenis sampel citra di mana yang satu citra yang bersumber dari internet dan yang satunya bersumber dari hasil foto kamera ponsel (Humonggio et al., 2019). Citra yang menjadi sampel adalah citra yang terdapat karakter teks di dalamnya. Di mana ada yang memiliki karakter teks dengan jumlah yang sedikit ada pula yang banyak, serta memiliki kerumitan yang berbeda. Program nantinya dapat

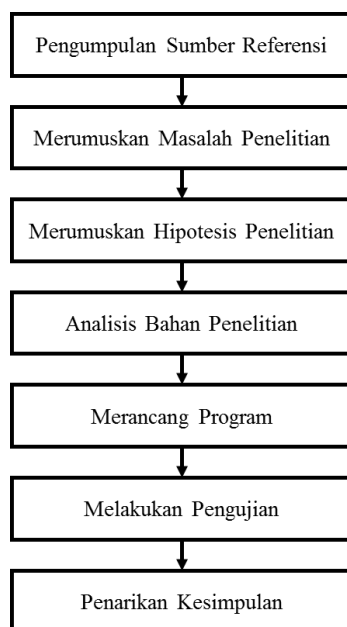
mengonversi citra yang menjadi inputan kemudian merubahnya menjadi teks sebagai outputnya

3.3 Teknik Pengumpulan Data

Pengumpulan data dilakukan peneliti dengan studi dokumen dimana dokumen yang menjadi data di pelajari sebagai sumber penunjang penelitian. Dimana pada penelitian ini menggunakan foto atau gambar sebagai objeknya sehingga data yang di peroleh berupa karakteristik dari gambar yang di dalamnya terdapat karakter teks seperti ukuran file gambar, jumlah karakter teks yang terdapat pada gambar, dan jumlah kata pada gambar (Duan, 2018).

3.4 Alur Pengolahan Data

Setelah mendapati bahan penelitian, kemudian dalam pengolahan data peneliti menjalani tahap tahap sebagai berikut :



Gambar 1. Tahap Penelitian

Dari bagan tersebut dapat di jelaskan mengenai tahap penelitian, berikut penjelasan dari bagan tersebut :

a. Pengumpulan Sumber Referensi

Tahap pertama sebelum melakukan penelitian terlebih dahulu peneliti mengumpulkan referensi yang memiliki hubungan dengan topik yang akan di teliti berupa karya ilmiah terdahulu dan yang bersumber dari buku. Pengumpulan referensi yang relevan dengan topik yang akan di ulas di penelitian, kemudian mempelajari metode yang di digunakan pada penelitian terdahulu.

b. Rumusan Masalah

Fokuskan topik pengujian nantinya, maka di perlukan rumusan masalah. Dengan membuat rumusan masalah maka peneliti mampu mengidentifikasi hubungan kesesuaian antara topik penelitian dengan bahan yang akan di pakai pada saat pengujian nantinya. Topik yang akan di ulas adalah bagaimana mengolah sebuah citra yang bersumber dari online maupun jepretan kamera yang di ada tulisanya dan diubah menjadi sebuah teks yang dapat di olah oleh komputer dengan aplikasi yang memanfaatkan metode *digital image processing*.

c. Merumuskan Hipotesis Penelitian

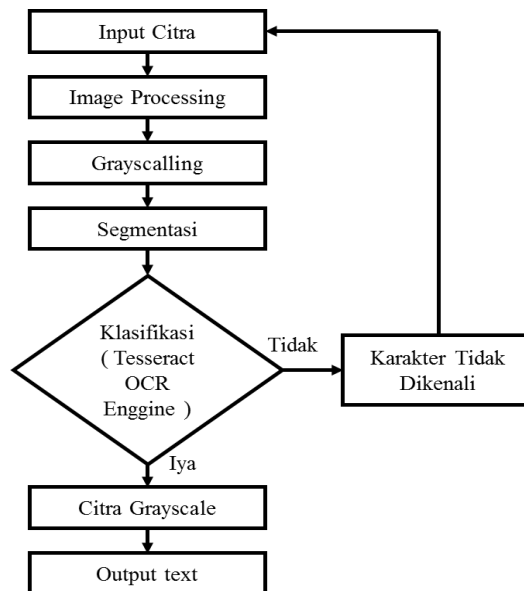
Hampir sama dengan perumusan masalah, tahap hipotesis juga di perlukan untuk mengidentifikasi kemungkinan kemungkinan yang dapat terjadi pada hasil penelitian nantinya. Hipotesis dibuat dengan sebuah variabel, dimana variabel tersebut memiliki kemungkinan terjadi maupun variabel yang minin kemungkinannya terjadi.

d. Analisisa Bahan Penelitian

Sebelum membuat sebuah program, peneliti mengumpulkan bahan bahan yang di perlukan. Dimana yang utamanya adalah citra yang akan di masukan dan dijadikan bahan penelitian, peneliti mengambil bahan citra yang bersumber dari internet dan kamera ponsel. Kemudian mengumpulkannya kedalam sebuah file di dalam komputer. Setelahnya peneliti menggunakan bahasa pemrograman python dengan menerapkan metode pengolahan citra digital. Sehingga perlu untuk menginstal fungsi fungsi yang di perlukan.

e. Perancangan Program

Tahap awal peneliti membuat program dalam bahasa program berbasis Python dan *running library* yang diperlukan. Beberapa sintaks pengkodean yang digunakan mempunyai karakteristik yang berkaitan dengan topik penelitian. Sebagai contoh pengenalan teks dari gambar, peneliti menggunakan model *Tesseract OCR* dan menambahkan sintaks pada program berbasis Python. (Ramadhani et al., 2019) Agar dapat memahami alur program, peneliti menggambarkannya sebagai diagram alir sistem dalam perancangan program pengolahan digital untuk pengenalan teks pada suatu gambar.



Gambar 2. Flowchart Program (Nasional et al., 2014)

Berdasarkan diagram alur prosedur pengoperasian dapat di jelaskan sebagai berikut :

1. Langkah pertama masukkan gambar ke dalam folder folder kerja dan beri nama dalam program yang dibuat di komputer menggunakan sintaks kode.
2. Langkah selanjutnya adalah pengolahan gambar, dimana kita dapat memastikan bahwa program mengenali gambar yang kita masukkan.
3. Kemudian proses pengolahan gambar yang dilakukan pada program tersebut adalah mengubah warna asli dari RGB menjadi greyscale. Tujuannya adalah untuk mengurangi informasi yang tidak perlu dalam gambar yang dapat dikenali.
4. Program kemudian memindai informasi yang diperoleh dari gambar dan memilih informasi yang diperlukan.

5. Kemudian, program mengenali teks pada gambar dengan memindai karakter gambar uji.
6. Program memindai dan menemukan atau tidak menemukan teks dalam gambar.
7. Hasil pengoperasian program menunjukkan gambar berwarna abu-abu yang menunjukkan keberhasilan pengolahan gambar digital yang dilakukan pada program.
8. Ketika teks terdeteksi pada gambar, program mengubahnya menjadi keluaran teks yang dapat diproses oleh komputer.

f. Cara kerja Tesseract dimulai dari tahap awal dilakukan Pra-Pemrosesan kemudian dilakukan Segmentasi Karakter, tahap selanjutnya Ekstraksi Ciri yang terakhir Pengenalan Karakter (Jawahar & Kumar, 2007).

g. Melakukan Pengujian

Setelah aplikasi dibuat dan bahan dalam hal ini gambar dimasukkan ke dalam file kerja program, langkah utama adalah pengujian. Pada tahap saat ini yang menjadi fokus keberhasilan penelitian, objek yang dikenali adalah gambar dengan teks, setelah itu program mengolahnya menjadi teks.

h. Penarikan Kesimpulan

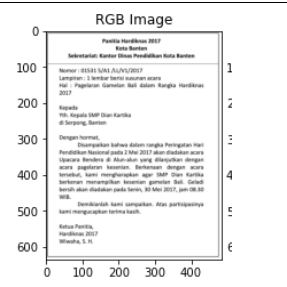
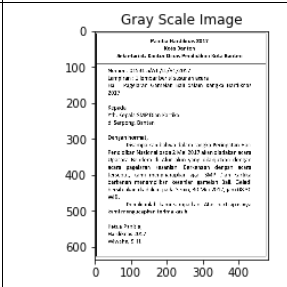
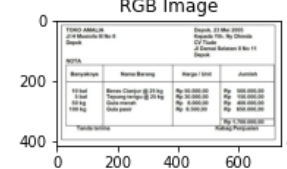
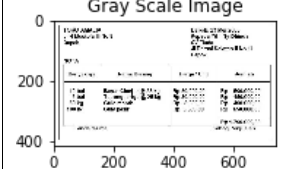
Program dapat dikatakan berhasil tatkala tahap tahap yang telah di rancang, di buat, serta dilakukan sebelumnya berjalan dan menghasilkan output yang sesuai dengan yang di harapkan. Perhitungan berfokus terhadap sebuah kata bukan terhadap karakter untuk mengetahui hasil yang lebih efektif untuk bisa di baca oleh manusia.



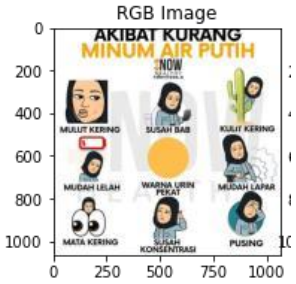


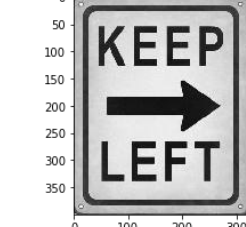
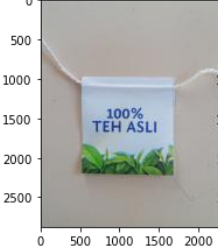
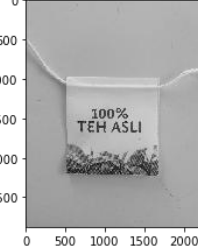

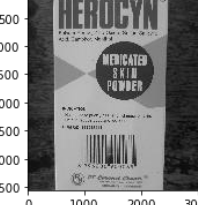


4. HASIL DAN PEMBAHASAN



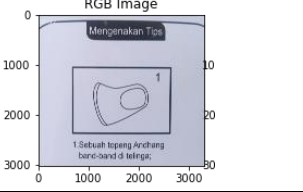
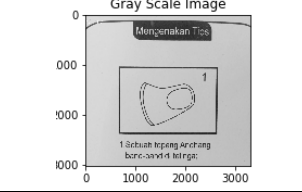
4.1 Tahap Pengolahan Citra Digital

- a. Sample citra berformat .jpg, gambar ini merupakan citra warna atau RGB. Kemudian mengumpulkannya ke dalam sebuah file direktori kerja program.
- b. Citra yang masih berupa citra warna tadi kemudian oleh program di ubah menjadi grayscale untuk membantu penekanan pada komposisi, pola, dan bentuk objek dalam gambar tanpa gangguan dari warna citra selanjutnya (Tian & Yin, 2018).

Tabel 1. RGB ke grayscale

| No | Nama File | Citra Warna / RGB | Citra Grayscale |
|----|---------------|---|--|
| 1 | Internet1.jpg |  |  |
| 2 | Internet2.jpg |  |  |

| No | Nama File | Citra Warna / RGB | Citra Grayscale |
|----|---------------|--|---|
| 3 | Internet3.jpg |  <p>RGB Image</p> |  <p>Gray Scale Image</p> |
| 4 | Internet4.jpg |  <p>RGB Image</p> |  <p>Gray Scale Image</p> |
| 5 | Internet5.jpg |  <p>RGB Image</p> |  <p>Gray Scale Image</p> |
| 6 | Kamera1.jpg |  <p>RGB Image</p> |  <p>Gray Scale Image</p> |
| 7 | Kamera2.jpg |  <p>RGB Image</p> |  <p>Gray Scale Image</p> |
| 8 | Kamera3.jpg |  <p>RGB Image</p> |  <p>Gray Scale Image</p> |

| No | Nama File | Citra Warna / RGB | Citra Grayscale |
|----|-------------|---|--|
| 9 | Kamera4.jpg |  |  |
| 10 | Kamera5.jpg |  |  |

c. Kemudian setelah menjadi grayscale program mengidentifikasi apakah terdapat karakter teks pada citra, jika di temukan karakter teks pada citra maka program akan mengonversikannya dan menghasilkan output teks.

4.2 Hasil Pengujian

Hasil dari program adalah untuk mengidentifikasi teks yang terdapat pada citra. Tets to voice 2020 WartawanPeneliti mengui 10 sample citra dengan tingkat kualitas gambar, kerumitan karakter, serta jumlah karakter teks yang berbeda beda sesuai yang di pilih peneliti yaitu:

- Karakter teks yang lebih dominan ketimbang objek yang lain.
- Karakter teks sedikit dengan objek lain yang lebih dominan
- Full karakter teks

Tabel 2. Sampel Citra

| No | Nama File Citra | Ukuran File (kb) | Jumlah Karakter Teks (Kata) |
|-----------|-----------------|-------------------|-------------------------------|
| 1 | Internet1.jpg | 32 | 105 |
| 2 | Internet2.jpg | 39 | 67 |
| 3 | Internet3.jpg | 101 | 16 |
| 4 | Internet4.jpg | 133 | 26 |
| 5 | Internet5.jpg | 24 | 2 |
| 6 | Kamera1.jpg | 1200 | 3 |
| 7 | Kamera2.jpg | 2500 | 39 |
| 8 | Kamera3.jpg | 2500 | 19 |
| 9 | Kamera4.jpg | 3000 | 9 |
| 10 | Kamera5.jpg | 1700 | 9 |
| Rata rata | | 1122,9 | 29,5 |

Dari Tabel 2 di dapat keterangan nama file citra yang di olah beserta ukuran filenya. Untuk mencari jumlah kata yang terdapat pada citra peneliti menggunakan perhitungan manual dengan excel, dengan menggunakan rumus pendekatan umum untuk menghitung jumlah kata pada sel excel yaitu:

$$JK=IF(LEN(TRIM(Q))=0,0,LEN(TRIM(Q))-LEN(SUBSTITUTE(Q," ", ""))+1)$$

(Fadjeri, Rahmawati, et al., 2022)

Keterangan :

- JK = Jumlah kata
- IF = Untuk membuat perbandingan logis antara nilai dengan perkiraan
- LEN = Untuk menghitung jumlah karakter dalam teks
- TRIM = Untuk menghapus karakter spasi dari teks kecuali spasi tunggal di antara kata
- Q = Teks yang diidentifikasi
- SUBSTITUTE = Untuk membuang karakter spasi dari teks

Dengan menggunakan 10 sample citra dengan ukuran yang berbeda-beda. Didapat hasil sebagai berikut:

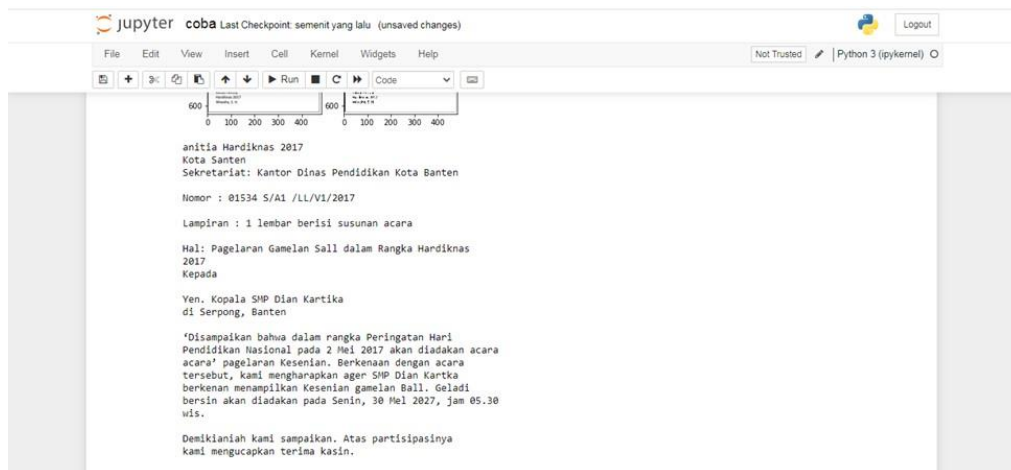
Tabel 3. Hasil Pengujian

| No | Nama File Citra | Jumlah Kata | Kata Yang Terdeteksi | Persentase |
|-----------|-----------------|-------------|----------------------|------------|
| 1 | Internet1.jpg | 105 | 80 | 76 % |
| 2 | Internet2.jpg | 67 | 58 | 86 % |
| 3 | Internet3.jpg | 16 | 16 | 100 % |
| 4 | Internet4.jpg | 26 | 11 | 42 % |
| 5 | Internet5.jpg | 2 | 2 | 100 % |
| 6 | Kamera1.jpg | 3 | 3 | 100 % |
| 7 | Kamera2.jpg | 39 | 15 | 38 % |
| 8 | Kamera3.jpg | 19 | 12 | 63 % |
| 9 | Kamera4.jpg | 9 | 5 | 55 % |
| 10 | Kamera5.jpg | 9 | 9 | 100 % |
| Rata rata | | 295 | 19,1 | 76 % |

Dari Tabel 3 dapat kita peroleh informasi jumlah kata yang berhasil diidentifikasi oleh program. Untuk menghitungnya masih sama dengan hasil program, peneliti menghitung jumlah kata dengan perhitungan manual di excel. Sementara itu, Anda juga dapat menggunakan rumus tersebut untuk menghitung tingkat keberhasilan program saat memindai teks dengan rumus berikut:

$$\text{Persentase} = (\text{Jumlah Kata Yang Terdeteksi} / \text{Jumlah Kata Keseluruhan}) \times 100 \%$$

(Fadjeri et al., 2016)



Gambar 3. Hasil Identifikasi Teks Pada Citra1

Dengan menggunakan file citra Internet1.jpg oleh program kemudian dapat teridentifikasi dan mendeteksi karakter teks sebanyak 80 kata dari 105 kata. Dari hasil yang terlihat pada Gambar 1, hasilnya berupa teks. Hasil dari akurasi aplikasi sebesar 76 % berhasil mengkonversikan citra ke dalam teks.

5. KESIMPULAN

Berdasarkan temuan penelitian dan pembahasan, maka dapat diambil kesimpulan dari penelitian ini sebagai berikut:

- Penelitian ini akan membuat suatu program pengolah citra digital yang akan menginput gambar yang mengandung karakter teks, mengenalinya dan langsung mencetaknya sebagai teks yang dapat diolah di komputer.
- Program ini dibuat menggunakan bahasa pemrograman Python. Dalam hal ini modul Tesseract digunakan untuk mengenali kata atau teks pada gambar masukan.
- Program ini menghasilkan akurasi 76% dari 10 sampel gambar. Dalam hal ini, sistem mengenali karakter teks pada gambar dengan akurasi rata-rata 81% yang berasal dari Internet, sedangkan akurasi data tangkapan kamera adalah 71%. Hasil uji akurasi menunjukkan bahwa program ini cukup akurat dan efektif dalam pengenalan karakter-gambar (I Komang Setia Buana, 2020).
- Modul Tesseract bahasa pemrograman Python dapat mendeteksi teks pada gambar dan mengubahnya langsung menjadi format teks yang dapat diproses oleh komputer, namun tidak semua gambar dapat dikenali 100% oleh program.
- Terdapat kata pada Gambar tersebut mengandung kata-kata yang tidak dikenali oleh program, sehingga keluaran teks yang dihasilkan tidak sesuai.

DAFTAR PUSTAKA

- Antara, A. M. E., Sari, S. A., Riswanti, N., Amin, D. A., Verdila, V., & Masa, A. P. A. (2023). Deteksi Nominal Rupiah Uang Kertas Berdasarkan Citra Warna Menggunakan Segmentasi K-Means Clustering dan Klasifikasi Random Forest. *Kreatif Teknologi Dan Sistem Informasi (KRETISI)*, 1(1), 34–39. <https://doi.org/10.30872/kretisi.v1i1.776>
- Arini, N. F., Ubaidillah, A., Wibisono, K. A., & Ulum, M. (2020). Identifikasi embrio dalam telur berbasis image processing. *Jurnal Teknik Elektro Dan Komputasi (ELKOM)*, 2(1), 11–19. <https://doi.org/10.32528/elkom.v2i1.3137>

- BURHAEIN, E., FADJERI, A., & WIDIYONO, I. P. (2023). Application of Naive Bayes Algorithm for Physical Fitness Level Classification. *International Journal of Disabilities Sports and Health Sciences*, 7(July 2023), 178–187. <https://doi.org/10.33438/ijdshs.1330745>
- Dan, R., Studi, A., Kopi, K., Arial, T., Fadjeri, A., Setyanto, A., & Kurniawan, M. P. (2020). *Pengolahan Citra Digital Untuk Menghitung Ekstraksi Ciri Greenbean Kopi*. 8(1), 8–13.
- Duan, Z. (2018). *Characters Recognition of Binary Image using KNN INTRODUCTION TO KNN*. 116–118.
- Effendi, M. M., Permana, A. Y., & Nawangsih, I. (2019). Penerapan Ekstraksi Image ke TXT dengan Optical Character Recognition untuk Otomatis Data Kependudukan. *Simposium Nasional Ilmiah Simponi*, November, 496–502. <https://doi.org/10.30998/simponi.v0i0.484>
- Fadjeri, A., Pd, S., & Kom, M. (n.d.). *Pengolahan Citra Digital Untuk Analisis Citra Jupyter Notebook Pengolahan Citra Digital Untuk Analisis Citra Jupyter Notebook*.
- Fadjeri, A., Rahmawati, A., & Fadilah, E. R. (2022). *Analisis Teks Bahasa Indonesia Dan Inggris Dari Sebuah Citra Menggunakan Pengolahan Citra Digital*. 10(2), 1–5.
- Fadjeri, A., Saputra, B. A., Kusuma, D., Ariyanto, A., & Kurniatin, L. (2022). *Karakteristik Morfologi Tanaman Selada Menggunakan Pengolahan Citra Digital*. 2, 1–12.
- Fadjeri, A., Studi, P., Matematika, P., & Purworejo, U. M. (2016). *PENGEMBANGAN MEDIA PEMBELAJARAN BERBASIS ICT (INFORMATION AND COMMUNICATION TECHNOLOGIES) PADA MATA PELAJARAN*. 103–108.
- Fernandes, D., Sunardi, H., & ZULKifli. (2023). Implementasi Deteksi Plat Nomor Kendaraan Bermotor Roda Dua Berbasis OpenCV Untuk Keamanan Parkir. *Prosiding Seminar Nasional Teknologi Komputer Dan Sains*, 1(1), 1–15.
- Humonggio, R., Abdullah, R. K., & Asri, M. (2019). Pengenalan Plat Nomor Menggunakan Image Processing Pada Perangkat Mikrokontroler. *Jurnal Teknologi Informasi Indonesia (JTII)*, 4(2), 63–70. <https://doi.org/10.30869/jtii.v4i2.400>
- I Komang Setia Buana. (2020). Implementasi Aplikasi Speech to Text untuk Memudahkan Wartawan Mencatat Wawancara dengan Python. *Jurnal Sistem Dan Informatika (JSI)*, 14(2), 135–142. <https://doi.org/10.30864/jsi.v14i2.293>
- Jawahar, C. V., & Kumar, A. (2007). Content-level annotation of large collection of printed document images. *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR*, 2, 799–803. <https://doi.org/10.1109/ICDAR.2007.4377025>
- Kembarawan, I. G. K. (n.d.). *MENITI TEKS AJARAN HINDU SEBAGAI INSPIRASI KARMA Abstrak SOPHIA DHARMA*, Vol 4 , No . 1 , Mei 2021 *Pendahuluan Teks karya sastra yang diwarisi oleh para pendahulu memiliki nilai yang sangat*. 36–60.
- Kurniawan, H., Sela, E. I., & Plate, L. (2023). *Indonesian Vehicle License Plate Identification Using YoloV5 and OCR*. 13(03), 913–920.
- Mohan, A., & Poobal, S. (2018). Crack detection using image processing: A critical review and analysis. *Alexandria Engineering Journal*, 57(2), 787–798. <https://doi.org/10.1016/j.aej.2017.01.020>
- Munantri, N. Z., Sofyan, H., & Florestiyanto, M. Y. (2020). Aplikasi Pengolahan Citra Digital Untuk Identifikasi Umur Pohon. *Telematika*, 16(2), 97. <https://doi.org/10.31315/telematika.v16i2.3183>
- Nabusa, Y. N. (2019). Pengolahan Citra Digital Perbandingan Metode Histogram Equalization Dan Spesification Pada Citra Abu-Abu. *J-Icon*, 7(1), 87–95.
- Nasional, S., Informasi, S., Nugroho, D. C., Sulistiyo, M. D., Purnama, B., Informatika, F.,

- & Telkom, U. (2014). *Optical Character Recognition Pada Smart Phone Menggunakan Contour Analysis Dan Feature*. September.
- Ramadhani, R. Z., Rusdianto, H., & Yahya, V. (2019). Rancang Bangun Aplikasi Pusat Informasi Sekolah Dengan Penerapan Chatbot Menggunakan Aiml Berbasis Android Pada Smk Otomotif Al Husna Tangerang. *JIKA (Jurnal Informatika)*, 3(2), 27–33. <https://doi.org/10.31000/jika.v3i2.2076>
- Rapar, P. V., Rorimpandey, G. C., Informatika, T., & Manado, U. N. (2023). Penerapan Teknologi Optical Character Recognition Pada Layanan Pendataan Pengunjung Di PT . Pertamina Geothermal Energy Area Lahendong. *INNOVATIVE: Journal Of Social Science Research*, 3(6), 6484–6496. 3
- Ripai, A., & Saputri, R. D. (2023). Efektivitas Penyimpanan Arsip Berbasis Digital Dalam Meningkatkan Keamanan Data Di Madrasah Tsanawiyah. *Jurnal Isema : Islamic Educational Management*, 8(2), 211–222. <https://doi.org/10.15575/isema.v8i2.25158>
- Rizky pratama, A. (2021). Klasifikasi Daging Sapi Berdasarkan Ciri Warna Dengan Metode Otsu dan K-Nearest Neighbor. *Techno Xplore : Jurnal Ilmu Komputer Dan Teknologi Informasi*, 6(1), 9–18. <https://doi.org/10.36805/technoxplore.v6i1.1239>
- Sugiarta, G., Andini, D. P., & Hidayatullah, S. (2021). Ekstraksi Informasi/Data e-KTP Menggunakan Optical Character Recognition Convolutional Neural Network. *JTERA (Jurnal Teknologi Rekayasa)*, 6(1), 1. <https://doi.org/10.31544/jtera.v6.i1.2021.1-6>
- Tian, J., & Yin, X. (2018). Adaptive image enhancement algorithm based on the model of surface roughness detection system. *Eurasip Journal on Image and Video Processing*, 2018(1), 1–9. <https://doi.org/10.1186/s13640-018-0343-1>
- Valentino, J., & Susetyo, Y. A. (2023). Analisis Perbandingan Optical Character Recognition Google Vision dengan Microsoft Computer Vision pada Pembacaan KTP-el. *Jurnal JTIK (Jurnal Teknologi Informasi Dan Komunikasi)*, 7(4), 552–561. <https://doi.org/10.35870/jtik.v7i4.1046>
- Y., K., I.G.W, S., & L. D. S., A. (2021). Upaya Meningkatkan Kemampuan Menulis Melalui Strategi Please. *Indonesian Gender and Society Journal*, 2(1), 10–17. <https://doi.org/10.23887/igsj.v2i1.39207>