



# GLOBAL THRESHOLDING IMPLEMENTATION FOR NOISE HANDLING IN DIGITAL IMAGE RECOGNITION

Dannu Purwanto<sup>1\*</sup>, Agustiyar<sup>2</sup>

<sup>1</sup> Program Studi Sains Data, Universitas Muhammadiyah Semarang

Jl. Kedungmundu No.18, Kedungmundu, Kota Semarang, Jawa Tengah Telp. (024) 76740296

<sup>2</sup> Program Studi Manajemen Informatika, AMIK JTC Semarang

Jl. Kelud Raya 19 Semarang

## ARTICLE INFO

### History of the article:

Received 6 februari 2024

Received in revised form 23 Februari 2024

Accepted 26 Februari 2024

Available online 28 Maret 2024

### Keywords:

Segmentation; Character Recognition; Global Thresholding; Noise

### \* Correspondence:

Telepon:

+62 85641273222

E-mail:

dannupurwanto@unimus.ac.id

## ABSTRACT

Text recognition (OCR - Optical Character Recognition) is a research field that is gaining widespread attention due to its wide application in

image and document processing. Although OCR technology has achieved a high level of success, the main challenge faced is the presence of noise in text image, noise causes decreased text recognition results, noise causes miss classification. Therefore needed noise handling text recognition. The aim of this research is to provide valuable insight into the techniques and approaches used in the context of noise treatment using global threshold methods. The method used starts from an input digital image, then preprocessing is carried out by converting the image into a gray scale image, then a threshold is applied to the image, then recognition is carried out. From 6 experiments, the best results were obtained for character recognition with a threshold value ( $t$ ) of 65 and a character recognition accuracy percentage of 94.29%.  $T$  value determined manually and static for separates the all object and the background, while in reality the lighting or contrast always varies. Suggestions for further research include developing an adaptive thresholding method approach to obtain threshold values automatically and optimally. So that if faced with varying lighting conditions or contrast, better results can be obtained.

## 1. INTRODUCTION

Pengenalan teks (OCR - *Optical Character Recognition*) merupakan bidang riset yang mendapatkan banyak perhatian karena aplikasinya yang luas dalam pemrosesan gambar dan dokumen. Sebagaimana yang dilakukan [1] dalam pengenalan kartu identitas pada proses *boarding* secara online, ekstraksi teks dari sebuah gambar [2], pengembangan teknik pengolahan citra untuk klasifikasi kendaraan [3], pencarian dokumen secara pintar [4], transkripsi siaran audio-video [5], *augmented reality navigation* [6]. Meskipun OCR banyak digunakan dalam pemrosesan gambar dan dokumen, akan tetapi pencahayaan yang tidak merata, buram, distorsi perspektif, orientasi, dan latar belakang yang kompleks menjadi tantangan pada bidang ini [7]. Kurangnya pencahayaan pada gambar menimbulkan tantangan besar terhadap pengenalan teks, karena menyebabkan variasi

warna, gaya font, serta kegelapan dan cahaya, sehingga sulit untuk dieksekusi dalam situasi real-time [8]. Selain itu, pengenalan teks semakin diperumit oleh variasi ukuran teks, warna, font, orientasi, latar belakang yang kompleks, oklusi, dan kondisi pencahayaan yang tidak merata [9]. Variabilitas yang tinggi dalam hal distorsi perspektif, tampilan fisik, pencahayaan variabel, dan desain jenis huruf dalam teks merupakan tantangan tersendiri [10]. Selain itu, efek iluminasi juga berpengaruh terhadap kinerja pengenalan teks dalam mendeteksi teks pada gambar [11]. Terlepas dari tantangan-tantangan ini, terdapat kemajuan yang signifikan dalam bidang deteksi dan pengenalan teks pada gambar pemandangan alam selama dekade terakhir [12]. Berbagai penelitian berfokus pada pengembangan model dan algoritma untuk mengatasi kompleksitas pengenalan teks termasuk pengembangan jaringan saraf tiruan yang dapat dilatih untuk pengenalan urutan berbasis gambar [13], deteksi teks pemandangan melalui proposal rotasi [7], dan algoritma berdasarkan koreksi posisi teks dan jaringan encoder-decoder untuk pengenalan teks sensor visual [14].

Teknologi OCR memiliki aplikasi yang luas dalam berbagai bidang. Salah satu contoh aplikasi dalam industri perbankan [15], di mana OCR digunakan untuk merekam dan mengekstrak data perbankan. Dalam dunia kesehatan [16], OCR digunakan untuk mendeteksi informasi obat. Di bidang transportasi [17], aplikasi OCR digunakan untuk membaca plat nomor kendaraan otomatis [18], dan memantau lalu lintas [19]. Tentunya dalam pengenalan Teks hal yang terpenting adalah akurasi dan efisiensi dalam pengenalan teks [20]. Akurasi dan efisiensi adalah dua aspek kritis dalam pengenalan teks melalui OCR. Akurasi yang tinggi diperlukan untuk memastikan bahwa teks yang dihasilkan dari gambar atau dokumen terbaca dengan benar dan tanpa kesalahan. Semakin tinggi akurasinya, semakin kecil kemungkinan sistem melakukan kesalahan saat mengenali teks. Hal ini penting terutama dalam konteks pengolahan informasi yang kritis, seperti dalam sistem perbankan, pengolahan dokumen hukum, atau rekam medis. Efisiensi juga menjadi faktor penting karena waktu proses yang cepat diperlukan dalam banyak aplikasi, terutama dalam skala besar di mana ribuan atau jutaan dokumen perlu diproses dalam waktu yang singkat. Kombinasi antara akurasi dan efisiensi memastikan bahwa teknologi OCR dapat diandalkan dan bermanfaat dalam skenario dunia nyata. Untuk meningkatkan akurasi tersebut teknologi OCR terus berkembang [21] [22].

Perkembangan terbaru dalam teknologi OCR telah melibatkan integrasi kecerdasan buatan (AI) [23]. Teknologi AI memungkinkan sistem OCR untuk belajar dan beradaptasi dengan pola-pola yang kompleks dalam teks dan gambar, sehingga meningkatkan tingkat akurasi secara signifikan. Penggunaan teknik deep learning [24], seperti jaringan saraf tiruan (neural networks), telah mengubah cara OCR bekerja. Model-model deep learning dapat mempelajari representasi yang lebih abstrak dari teks dan gambar, mengatasi tantangan yang kompleks seperti variasi font, orientasi, dan kondisi pencahayaan yang tidak merata. Selain itu, integrasi AI juga memungkinkan sistem OCR untuk melakukan tugas-tugas tambahan, seperti segmentasi gambar, pemrosesan bahasa alami, dan klasifikasi dokumen. Ini membuka peluang baru untuk aplikasi yang lebih canggih dan kompleks dalam berbagai industri, mulai dari layanan perbankan digital hingga pengelolaan data medis. Dengan terus berkembangnya teknologi AI, diharapkan bahwa kemampuan OCR akan terus meningkat, menjadikannya alat yang lebih efektif dan efisien dalam pemrosesan teks dan gambar dalam berbagai konteks aplikasi.

Teknologi OCR telah mencapai tingkat keberhasilan yang tinggi, namun tantangan utama yang dihadapi adalah adanya noise pada gambar teks [25]. Noise dapat berasal dari berbagai sumber, seperti pencacahan, kualitas gambar rendah, atau kekurangan pencahayaan, dan dapat menghambat akurasi sistem pengenalan teks. Salah satu pendekatan yang telah diterapkan untuk mengatasi noise pada gambar teks adalah menggunakan teknik thresholding [26]. Teknik ini melibatkan pemilihan nilai batas tertentu yang memisahkan antara piksel yang mengandung informasi teks dan piksel yang tergolong sebagai noise. Meskipun teknik ini sederhana,

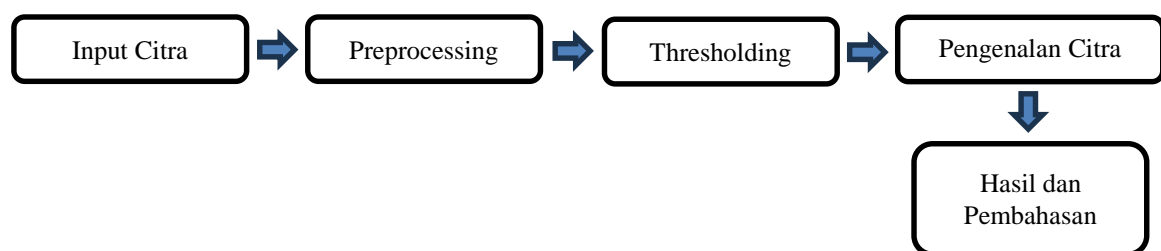
keberhasilannya sangat tergantung pada pemilihan nilai threshold dan kemampuannya untuk menangani variasi noise yang kompleks.

Metode global Thresholding merupakan teknik fundamental dalam mengoptimalkan pemisahan objek dan background dalam segmentasi citra [27]. Ini melibatkan identifikasi nilai ambang batas yang membedakan antara piksel objek dan latar belakang, sehingga membantu dalam proses segmentasi. Dalam sebuah penelitian yang membandingkan metode segmentasi untuk pembuluh darah retina, tiga teknik ambang batas, yaitu Adaptive Thresholding, Binary Thresholding, dan Otsu Thresholding digunakan dalam proses segmentasi yang menyoroti pentingnya ambang batas dalam analisis gambar [28]. Selain itu, dalam konteks pemisahan sel kanker payudara, tahap awal melibatkan identifikasi sel kanker berdasarkan metode Global Thresholding, yang semakin menekankan perannya dalam analisis citra biomedis [29]. Meskipun beberapa hasil positif telah dicapai, masih ada kebutuhan untuk meningkatkan ketangguhan dan generalitas pendekatan ini terhadap variasi jenis noise. Implementasi teknik thresholding dalam pengenalan teks dapat memberikan kontribusi yang signifikan dalam meningkatkan akurasi dan keandalan sistem pengenalan teks, terutama dalam menghadapi noise. Penelitian ini bertujuan mengatasi noise pada OCR dengan metode global thresholding.

Metode Global thresholding merupakan metode yang mengkonversi seluruh piksel pada citra menjadi hitam dan putih dengan satu nilai ambang (threshold) digunakan untuk memisahkan piksel menjadi dua kelas: objek dan latar belakang [30]. Nilai ambang ini diterapkan secara seragam ke seluruh citra. Pada beberapa kasus, nilai ambang ( $t$ ) ini dapat ditentukan secara manual oleh pengguna, sesuai dengan karakteristik citra yang sedang diolah.

## 2. RESEARCH METHODS

Penelitian ini mengikuti diagram alir penelitian seperti ditunjukkan pada Gambar 1. Pada citra digital dengan model warna RGB (Red, Green, Blue) dengan ekstensi file JPG (Joint Photographic Experts Group).



Gambar 1. Diagram Alir Penelitian

### 2.1 Input Citra

Penelitian ini menggunakan citra digital dengan model warna RGB (*Red, Green, Blue*), dan berekstensi JPG (Joint Photographic Experts Group) yang di dalamnya terdapat noise. Gambar 2a menunjukkan citra input yang digunakan pada penelitian ini dengan model warna RGB. Kemudian gambar 2b adalah citra yang sudah dirubah menjadi *grayscale* yang kemudian akan dilakukan tahapan *preprocessing*.

### 2.2 Preprocessing

Pada tahapan ini citra RGB di konversi ke dalam model warna citra grayscale. Sehingga hanya memiliki satu nilai keabuan dalam tiap piksel. Proses perubahan citra RGB ke grayscale ditunjukkan pada Gambar 2b.

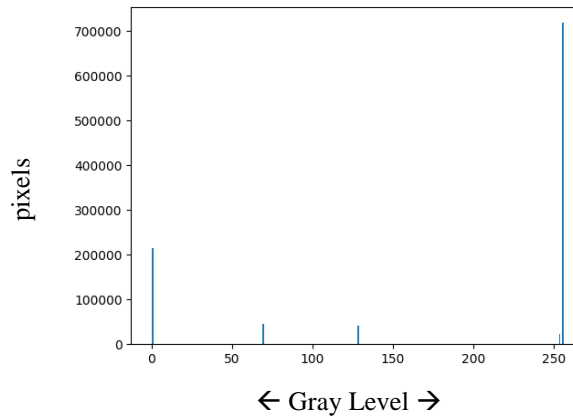


(a) Citra RGB

(b) Citra *Grayscale*Gambar 2. Perubahan Citra RGB ke *Grayscale*

### 2.3 Proses Segmentasi Citra

Sebelum proses segmentasi citra dilakukan akan ditentukan nilai ambang ( $t$ ) dengan metode *thresholding global* yang akan digunakan untuk memisahkan objek dan latar belakang. Nilai ini dipilih secara manual berdasarkan pemahaman tentang distribusi intensitas piksel pada citra. Iterasi dilakukan melalui setiap piksel pada citra kemudian membandingkan intensitas piksel dengan nilai ambang ( $t$ ). Kemudian akan dilakukan pelabelan pada piksel sebagai bagian dari objek atau latar belakang berdasarkan hasil perbandingan. Gambar 3 merupakan histogram dari citra *grayscale* yang dipakai. Dari histogram gambar 3 dilakukan pengamatan untuk iterasi nilai ambang ( $t$ ). Iterasi akan dilakukan sampai menemukan nilai ambang ( $t$ ) yang sesuai dan karakter dalam image dapat dikenali dan memperoleh persentase ketepatan yang terbaik.

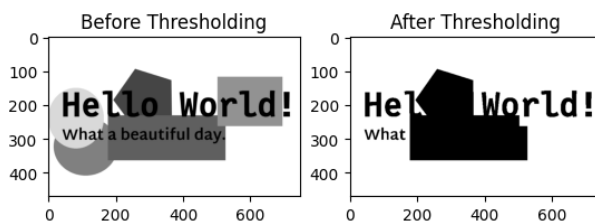


Gambar 3. Histogram Citra *Grayscale*

### 3. RESULTS AND DISCUSSION

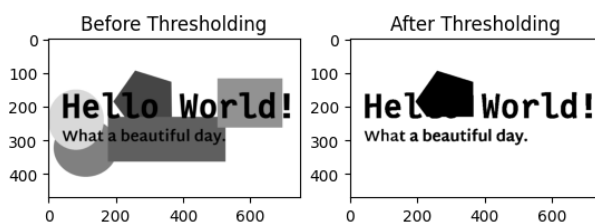
Dalam iterasi yang dilakukan, nilai ambang dipilih dengan mengamati image histogram pada image citra yang digunakan. Nilai ini dipilih secara manual berdasarkan pemahaman tentang distribusi intensitas piksel pada citra. Hal ini dilakukan untuk memisahkan antara objek dan background. Dari gambar 3 nilai ambang 100 di gunakan menjadi acuan awal karena dianggap dapat mewakili titik awal objek dan background. Nilai  $t$  yang di pilih manual kemudian dilakukan pengenalan dan dihitung persentase kecepatannya menggunakan similarity rasio. Secara matematis similarity rasio dihitung sebagai jumlah elemen yang cocok (*matching elements*) dibagi dengan jumlah elemen total (*total elements*) dalam dua urutan sampai karakter yang ada di dalam image terbaca dengan baik. Dalam penelitian ini ditampilkan 6 kali iterasi dengan nilai  $t$  yang berbeda. Karakter yang ingin dikenali dalam image adalah “Hello World! What a beautiful day.”

#### Percobaan 1



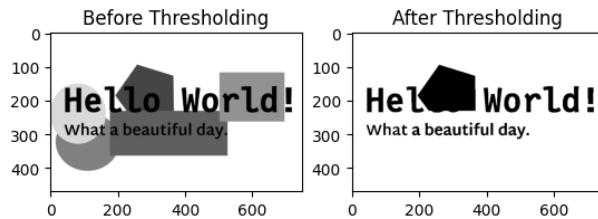
Percobaan pertama dilakukan dengan nilai  $t = 100$  pada image yang telah dilakukan grayscale. Setelah dilakukan pengenalan karakter yang dihasilkan adalah “He What” dan persentase ketepatan: 27.27%.

#### Percobaan 2



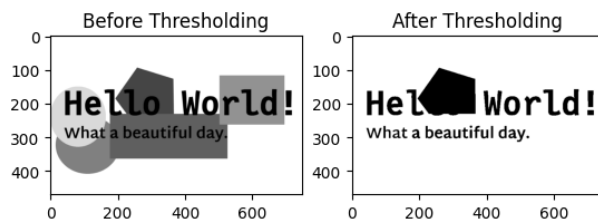
Pada percobaan kedua, dilakukan pengenalan karakter dengan nilai  $t = 90$  menghasilkan “He1@MI world! What a beautiful day.” dan persentase ketepatan: 87.32%.

## Percobaan 3



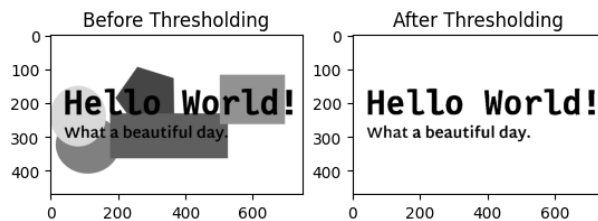
Pada percobaan ketiga, dilakukan pengenalan karakter dengan nilai  $t = 80$  menghasilkan “He@MI world! What a beautiful day.” dan Persentase Ketepatan: 88.57%.

## Percobaan 4



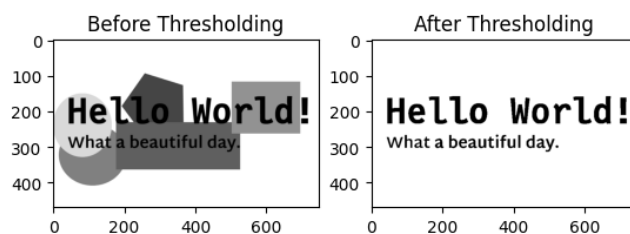
Percobaan keempat, dilakukan pengenalan karakter dengan nilai  $t = 70$  menghasilkan “Hel World! What a beautiful day.” dan persentase ketepatan: 91.18%.

## Percobaan 5



Percobaan kelima dilakukan dengan nilai  $t = 60$  pada image yang telah dilakukan grayscale. Setelah dilakukan pengenalan karakter, yang terbaca adalah “HeLlo World! What a beautiful day.” dan persentase ketepatan: 94.29%.

## Percobaan 6



Percobaan kelima dilakukan dengan nilai  $t = 65$  pada image yang telah dilakukan grayscale. Setelah dilakukan pengenalan karakter, yang terbaca adalah “Hello World! What a beautiful day.” dan persentase ketepatan: 94.29%.

#### 4. CONCLUSIONS AND RECOMMENDATIONS

Dari enam percobaan yang telah dilakukan, percobaan ke-6 dengan nilai  $t = 65$  mempunyai hasil terbaik dan mendapatkan persentase ketepatan pengenalan karakter sebesar 94.29% serta hasil karakter yang di kenali sama dengan karakter sebenarnya yaitu "Hello World! What a beautiful day." Metode global tresholding masih memiliki kelemahan di mana nilai  $t$  ditentukan secara manual, sehingga membutuhkan banyak percobaan. Serta nilai  $t$  statis di mana satu nilai ambang digunakan untuk seluruh citra dalam memisahkan objek dan latar belakang, sedangkan pada kenyataannya pencahayaan atau kontras selalu bervariasi. Hal ini berakibat beberapa karakter tidak dapat terdeteksi dengan benar. Saran untuk penelitian selanjutnya dapat dikembangkan pendekatan metode adaptive tresholding untuk memperoleh nilai ambang secara otomatis dan optimal. Sehingga jika dihadapkan dengan kondisi pencahayaan atau kontras yang bervariasi dapat diperoleh hasil yang lebih baik.

#### REFERENCES

- [1] B. Gulnara and A. Yerassyl, "Using Image Processing and Optical Character Recognition to Recognise ID cards in the Online Process of Onboarding," *SIST 2022 - 2022 Int. Conf. Smart Inf. Syst. Technol. Proc.*, 2022.
- [2] C. Kaundilya, D. Chawla, and Y. Chopra, "Automated text extraction from images using OCR system," *Proc. 2019 6th Int. Conf. Comput. Sustain. Glob. Dev. INDIACom 2019*, pp. 145–150, 2019.
- [3] I. S. In and I. S. In, "Development of an Image Processing Techniques for Vehicle Classification Using OCR and SVM."
- [4] A. M. Abbas, M. S. S. Hameed, S. Balakrishnan, and K. S. Anandh, "Intelligent Document Finding using Optical Character Recognition and Tagging," *Int. Conf. Autom. Comput. Renew. Syst. ICACRS 2022 - Proc.*, pp. 1165–1168, 2022.
- [5] J. Chaloupka, K. Palecek, P. Cerva, and J. Zdansky, "Optical character recognition for audio-visual broadcast transcription system," *11th IEEE Int. Conf. Cogn. Infocommunications, CogInfoCom 2020 - Proc.*, pp. 229–232, 2020.
- [6] I. Pivavaruk and J. R. Fonseca Cacho, "OCR Enhanced Augmented Reality Indoor Navigation," *Proc. - 2022 IEEE Int. Conf. Artif. Intell. Virtual Reality, AIVR 2022*, pp. 186–192, 2022.
- [7] J. Ma *et al.*, "Arbitrary-oriented scene text detection via rotation proposals," *IEEE Trans. Multimed.*, vol. 20, no. 11, pp. 3111–3122, 2018.
- [8] S. K. Devi and C. N. Subalalitha, "Intelligent Deep Learning Empowered Text Detection Model from Natural Scene Images," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 12, no. 3, pp. 1263–1268, 2022.
- [9] A. A. Chandio, M. Asikuzzaman, M. R. Pickering, and M. Leghari, "Cursive Text Recognition in Natural Scene Images Using Deep Convolutional Recurrent Neural Network," *IEEE Access*, vol. 10, pp. 10062–10078, 2022.
- [10] L. Gomez, A. Nicolaou, and D. Karatzas, "Improving patch-based scene text script identification with ensembles of conjoined networks," *Pattern Recognit.*, vol. 67, pp. 85–96, 2017.
- [11] T. Titijaronroj and K. Woraratpanya, "Iteration-free Bi-dimensional empirical mode decomposition and its application," *IEICE Trans. Inf. Syst.*, vol. E100D, no. 9, pp. 2183–2196, 2017.
- [12] H. Mokayed, P. Shivakumara, R. Saini, M. Liwicki, L. Chee Hin, and U. Pal, "Anomaly Detection in Natural Scene Images Based on Enhanced Fine-Grained Saliency and Fuzzy Logic," *IEEE Access*, vol. 9, pp. 129102–129109, 2021.
- [13] C. L. Cameras, O. Enqvist, and F. Kahl, "An End-to-End Trainable Neural Network for

- Image-Based Sequence Recognition and Its Application to Scene Text Recognition,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 7, pp. 1455–1461, 2017.
- [14] Z. Huang *et al.*, “An algorithm based on text position correction and encoder-decoder network for text recognition in the scene image of visual sensors,” *Sensors (Switzerland)*, vol. 20, no. 10, 2020.
- [15] M. K. Sharma, T. Sharma, M. Gupta, and Shivanshdeep, “Recognition and Extraction of Banking Data using OCR,” *Proc. - 2021 3rd Int. Conf. Adv. Comput. Commun. Control Networking, ICAC3N 2021*, pp. 472–475, 2021.
- [16] S. Godbole, D. Jojode, K. Kadam, and S. Karoshi, “Detection of Medicine Information with Optical Character Recognition using Android,” *Proc. B-HTC 2020 - 1st IEEE Bangalore Humanit. Technol. Conf.*, 2020.
- [17] I. S. In and I. S. In, “An Automated Vehicle Tracking System Using Haar-Cascade Classifiers and Optical Character Recognition Engine.”
- [18] D. Joshi and N. Mohd, “Techniques used in Automatic Number Plate Recognition,” *2023 4th Int. Conf. Emerg. Technol. INCET 2023*, 2023.
- [19] F. Chen, H. Kataoka, and Y. Satoh, “Text Detection in Traffic Informatory Signs Using Synthetic Data,” *Proc. Int. Conf. Doc. Anal. Recognition, ICDAR*, vol. 1, pp. 851–858, 2017.
- [20] M. Krishnamoorthi, K. P. S. Ram, M. Sathyan, and T. Vasanth, “Improving Optical Character Recognition(OCR) Accuracy using Multi-Layer Perceptron(MLP),” *7th Int. Conf. Trends Electron. Informatics, ICOEI 2023 - Proc.*, pp. 1642–1647, 2023.
- [21] P. Sisodia, S. Wajahat, and A. Rizvi, “Optical Character Recognition Development Using Python,” vol. 04, no. 075, pp. 1–13, 2023.
- [22] S. J. Jang, “Ocr related technology trends,” vol. 8, no. 1, pp. 13–20, 2020.
- [23] I. S. In and I. S. In, “Robotic Process Automation with AI and OCR to Improve Business Process : Review.”
- [24] M. Namysl and I. Konya, “Efficient, lexicon-free OCR using deep learning,” *Proc. Int. Conf. Doc. Anal. Recognition, ICDAR*, pp. 295–301, 2019.
- [25] K. I. Gunawan and J. Santoso, “Multilabel Text Classification Menggunakan SVM dan Doc2Vec Classification Pada Dokumen Berita Bahasa Indonesia,” *J. Inf. Syst. Hosp. Technol.*, vol. 3, no. 01, pp. 29–38, 2021.
- [26] Q. Mazhar, A. Masood, I. Touqir, and A. Ahmad, “Image De-Noising and Compression Using Statistical based Thresholding in 2-D Discrete Wavelet Transform,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 7, no. 11, pp. 311–316, 2016.
- [27] W. Gunawan and A. Z. Arifin, “Lokal Fuzzy Thresholding Berdasarkan Pengukuran Fuzzy Similarity Pada Interaktif Segmentasi Citra Panoramik Gigi,” *J. Infotel*, vol. 9, no. 1, p. 40, 2017.
- [28] A. Desiani, D. A. Zayanti, R. Primartha, F. Efriliyanti, and N. A. C. Andriani, “Variasi Thresholding untuk Segmentasi Pembuluh Darah Citra Retina,” *J. Edukasi dan Penelit. Inform.*, vol. 7, no. 2, p. 255, 2021.
- [29] D. Tuwohingide and C. Faticah, “Spatial Fuzzy C-means dan Rapid Region Merging untuk Pemisahan Sel Kanker Payudara,” *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 6, no. 1, 2017.
- [30] D. P. Agyztia Premana, Raden Mohamad Herdian Bhakti, “Segmentasi K-Means Clustering Pada Citra Menggunakan Ekstrasi Fitur Warna dan Tekstur,” *Pap. Knowl. . Towar. a Media Hist. Doc.*, 2020.