ISSN: 2598-8565 (media cetak)
ISSN: 2620-4339 (media online)

ANALISIS PERBANDINGAN *THRESHOLDING OTSU* DAN *SAUVOLA* UNTUK BINARISASI DOKUMEN SEMI-TERSTRUKTUR

[1,2]Zulfikar Yusya Mubarak✉, [1]Imam Riadi, [1]Abdul Fadlil
[1]Universitas Ahmad Dahlan, Yogyakarta, Indonesia
[2]Universitas Al-Irsyad Cilacap, Cilacap, Indonesia
Email: zulfikaryusya@gmail.com

DOI: https://doi.org/10.46880/jmika.Vol10No2.pp691-700

ABSTRACT

Semi-structured transaction documents acquired by scanners and cameras exhibit different degradation patterns. This study evaluates Otsu and Sauvola as baseline binarization methods on 100 self-collected document images, balanced between scanner and camera acquisitions (50 images each). Each image was converted to grayscale, enhanced with CLAHE, processed using six limited parameter configurations, and compared with manually validated pixel-level binary ground truth. Performance was assessed using MSE, PSNR, SSIM, Precision, Recall, F1-score, Foreground Ratio, and DRD, with F1-score as the primary criterion. Otsu_T10 achieved the best scanner-domain performance (F1-score = 0.917; DRD = 0.166). For camera images, Otsu_T11 produced the highest F1-score (0.172), while Sauvola_k03 yielded the lowest DRD (0.136). Visual failure analysis showed that scanner errors were dominated by broken thin strokes or character thickening, whereas camera errors included disconnected characters, missing small digits and table lines, and thick foreground artifacts caused by shadows and dark borders. These results confirm that the tested methods are baseline approaches rather than a complete camera-document processing solution. Perspective-aware cropping or correction, illumination normalization, and shadow reduction are mandatory before thresholding and downstream OCR on camera images. The conclusions are limited to the tested dataset and parameter ranges.

Keyword: Document Binarization, Otsu Thresholding, Sauvola Thresholding, Semi-Structured Documents, MSMEs.

ABSTRAK

Dokumen transaksi semi-terstruktur yang diperoleh melalui pemindai dan kamera menunjukkan pola degradasi berbeda. Penelitian ini mengevaluasi Otsu dan Sauvola sebagai metode binarisasi dasar pada 100 citra yang dikumpulkan mandiri, terdiri atas 50 citra pemindai dan 50 citra kamera. Setiap citra dikonversi ke skala abu-abu, ditingkatkan menggunakan CLAHE, diproses dengan enam konfigurasi, dan dibandingkan dengan ground truth biner tingkat piksel yang divalidasi manual. Kinerja diukur menggunakan MSE, PSNR, SSIM, Precision, Recall, F1-score, Foreground Ratio, dan DRD, dengan F1-score sebagai kriteria utama. Pada domain pemindai, Otsu_T10 memberikan hasil terbaik dengan F1-score 0,917 dan DRD 0,166. Pada domain kamera, Otsu_T11 memperoleh F1-score tertinggi sebesar 0,172, sedangkan Sauvola_k03 menghasilkan DRD terendah sebesar 0,136. Analisis visual menunjukkan bahwa kesalahan pada citra pemindai terutama berupa goresan tipis terputus dan penebalan karakter, sedangkan citra kamera mengalami karakter terputus, hilangnya angka kecil dan garis tabel, serta artefak foreground akibat bayangan dan tepian gelap. Hasil ini menegaskan bahwa kedua metode masih merupakan baseline. Citra kamera memerlukan pemotongan atau koreksi perspektif, normalisasi iluminasi, dan reduksi bayangan sebelum thresholding dan OCR. Temuan terbatas pada dataset dan rentang parameter yang diuji.

Kata Kunci: Binarisasi Dokumen, Otsu Thresholding, Sauvola Thresholding, Dokumen Semi Terstruktur, UMKM.

PENDAHULUAN

Digitalisasi dokumen merupakan bagian penting dalam pengelolaan informasi karena dokumen fisik perlu dikonversi ke dalam bentuk digital agar lebih mudah disimpan, diproses, ditelusuri, dan dianalisis (Huang et al., 2019; Feng et al., 2021; Ha & Horak, 2022). Proses akuisisi dokumen dapat dilakukan menggunakan *scanner* maupun kamera. *Scanner* menghasilkan citra dengan pencahayaan yang relatif stabil dan distorsi geometris yang rendah. Sebaliknya, kamera lebih praktis digunakan, tetapi citra yang dihasilkan lebih bervariasi karena dipengaruhi oleh

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

pencahayaan, bayangan, fokus, resolusi, dan sudut pengambilan gambar (Abbas & Hussain, 2017; Das et al., 2019).

Dalam aktivitas transaksi UMKM, dokumen seperti nota, faktur, bukti pembelian, surat jalan, dan catatan penjualan sering memiliki format semi-terstruktur. Dokumen tersebut memuat kombinasi nama barang, jumlah, harga satuan, harga total, tanggal, garis tabel, tanda tangan, dan ruang kosong. Sebelum catatan tersebut dapat mendukung pembukuan digital, entri data berbasis OCR (*Optical Character Recognition)*, pengarsipan transaksi, atau pelaporan usaha, citra dokumen harus melalui tahap praproses yang andal. Binarisasi merupakan salah satu tahap terpenting karena memisahkan informasi transaksi dari latar belakang dokumen (Li & Yu, 2025).

Perbedaan karakteristik akuisisi secara langsung memengaruhi tahap praproses citra dokumen. Salah satu tahap penting dalam praproses adalah binarisasi (Yang et al., 2024; Quattrini et al., 2024; Xiong et al., 2021; Bella et al., 2022; Han et al., 2023; Zhang et al., 2024). Binarisasi mengubah citra berwarna atau citra skala abu-abu menjadi citra biner dengan memisahkan objek utama dari latar belakang. Pada dokumen semi-terstruktur, objek utama tidak hanya berupa teks, tetapi juga angka, garis tabel, kolom, simbol, dan elemen informasi lainnya. Apabila proses binarisasi tidak optimal, karakter dapat terpotong, angka dapat hilang, garis tabel dapat rusak, dan derau latar belakang dapat terbaca sebagai foreground (Gatos et al., 2006).

Permasalahan ini bersifat kritis pada dokumen transaksi UMKM karena kesalahan kecil pada objek utamanya *(foreground)* dapat mengubah makna catatan transaksi. Sebagai contoh, hilangnya digit dapat mengubah nilai nominal, garis tabel yang terputus dapat mengganggu penguraian dokumen, dan bayangan dapat ditafsirkan sebagai informasi transaksi yang keliru. Oleh karena itu, metode binarisasi yang dipilih tidak hanya harus menghasilkan tampilan visual yang bersih, tetapi juga harus mempertahankan objek *foreground* yang berkaitan dengan transaksi secara akurat (Thendral & Abitha, 2026; Kasem et al., 2026).

Tantangan menjadi lebih kompleks ketika dokumen berasal dari domain akuisisi yang berbeda. Citra *scanner* umumnya dipengaruhi oleh kualitas cetak, noda, atau ketebalan karakter, sedangkan citra kamera lebih sering dipengaruhi oleh pencahayaan tidak merata, bayangan, blur, kontras rendah, dan distorsi perspektif. Akibatnya, satu konfigurasi thresholding belum tentu memberikan kinerja yang setara pada kedua domain tersebut (Darpito et al., 2026; Pratikakis et al., 2017).

Penelitian ini membandingkan empat parameter utama, yaitu metode *thresholding* berupa Otsu dan Sauvola; domain akuisisi berupa *scanner* dan kamera; parameter konfigurasi berupa Otsu_T09, Otsu_T10, Otsu_T11, Sauvola_k03, Sauvola_k04, dan Sauvola_k05; serta metrik evaluasi berupa kualitas visual dan akurasi segmentasi. Penentuan metode terbaik tidak hanya didasarkan pada tampilan visual. Standar keputusan utama adalah *F1-score* karena metrik ini menyeimbangkan *Precision* dan *Recall*.

Precision menunjukkan kemampuan untuk mencegah derau latar belakang diklasifikasikan sebagai informasi transaksi, sedangkan *Recall* menunjukkan kemampuan mempertahankan elemen transaksi penting, seperti teks, angka, nilai nominal, dan garis tabel. DRD (*Distance Reciprocal Distortion*) digunakan sebagai standar pelengkap untuk menilai distorsi visual, sedangkan MSE *(Mean Squared Error)*, PSNR *(Peak Signal-to-Noise Ratio)*, dan SSIM *(Structural Similarity Index Measure)* digunakan untuk menilai kemiripan visual.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan mengevaluasi kinerja *thresholding* Otsu dan *thresholding* Sauvola dalam proses binarisasi dokumen transaksi UMKM semi-terstruktur yang diperoleh melalui akuisisi *scanner* dan kamera. Evaluasi difokuskan pada kemampuan setiap metode dalam memisahkan informasi *foreground* dari latar belakang, mempertahankan keterbacaan teks dan angka, menjaga struktur dokumen seperti garis, tabel, dan kolom, serta mengurangi distorsi visual yang dapat memengaruhi proses OCR dan pengolahan data transaksi pada tahap berikutnya

TINJAUAN PUSTAKA

Binarisasi citra dokumen memisahkan informasi *foreground* dari latar belakang dan tetap menjadi tahap praproses mendasar dalam OCR serta analisis dokumen. Kajian terkini mengelompokkan pendekatan utama ke dalam teknik global, lokal-adaptif, berbasis pembelajaran, dan hibrida, sekaligus menekankan tantangan yang masih berlangsung pada dokumen terdegradasi dan dokumen dengan pencahayaan tidak merata (Yang et al., 2024; Quattrini et al., 2024). Metode mendalam dan hibrida dapat meningkatkan ketahanan, tetapi kebutuhan komputasi dan datanya masih lebih tinggi dibandingkan metode *thresholding* klasik (Xiong et al., 2021; Bella et al., 2022; Han et al., 2023; Zhang et al., 2024).

Thresholding Otsu merupakan pendekatan global yang memilih satu nilai ambang dari histogram skala abu-abu. Kesederhanaan dan efisiensi komputasinya menjadikan metode ini menarik ketika

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

ISSN: 2598-8565 (media cetak)
ISSN: 2620-4339 (media online)

intensitas *foreground* dan latar belakang terpisah dengan baik, tetapi kinerjanya dapat menurun pada variasi pencahayaan lokal (Jing & Tang, 2024). *Thresholding* Sauvola menentukan nilai ambang berdasarkan nilai rata-rata dan simpangan baku lokal sehingga lebih adaptif terhadap latar belakang yang tidak seragam; namun, hasilnya tetap sensitif terhadap ukuran jendela lokal dan parameter sensitivitas k (Li et al., 2021; Bloechle et al., 2023).

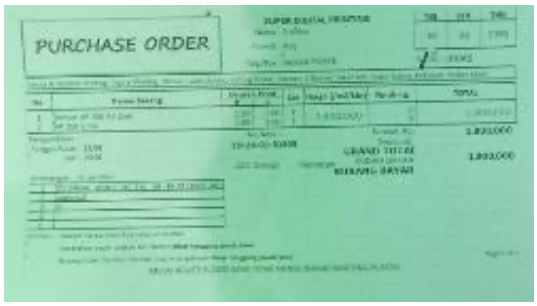
Dokumen yang ditangkap menggunakan kamera menimbulkan gangguan geometris dan fotometris yang tidak terlalu menonjol pada citra *scanner*. Penelitian mengenai pemulihan homografi, dewarping, dan koreksi pencahayaan menunjukkan bahwa sudut pengambilan, perspektif, dan pencahayaan lokal secara langsung memengaruhi keterbacaan dokumen serta pengenalan pada tahap berikutnya (Abbas & Hussain, 2017; Das et al., 2019; Feng et al., 2021). Oleh karena itu, penelitian ini mengevaluasi konfigurasi Otsu dan Sauvola klasik secara terpisah pada domain *scanner* dan kamera dengan menggunakan metrik kualitas visual dan segmentasi *foreground*.

METODE PENELITIAN

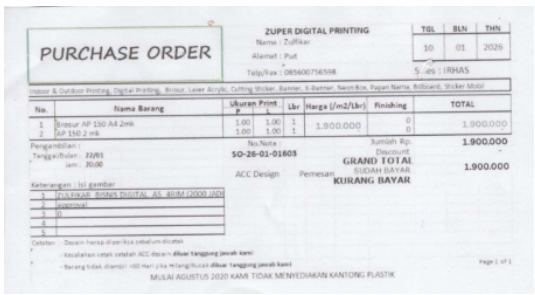
Penelitian ini menggunakan desain eksperimen komparatif untuk mengevaluasi *thresholding* Otsu dan Sauvola pada citra dokumen semi-terstruktur yang diperoleh dari *scanner* dan kamera (Darpito et al., 2026). Data terdiri atas dokumen transaksi yang memuat teks, angka, garis tabel, kolom, area kosong, dan nilai nominal. *Ground truth* biner digunakan sebagai acuan untuk mempertahankan *foreground*. Citra yang tidak lengkap, terhalang, atau sejak awal tidak terbaca dikeluarkan agar evaluasi berfokus pada kinerja binarisasi.

Dataset penelitian terdiri atas 100 citra dokumen transaksi semi-terstruktur yang dikumpulkan secara mandiri dan dibagi seimbang menjadi 50 citra *scanner* dan 50 citra kamera. Pembagian 50:50 digunakan agar perbedaan kinerja dapat dikaitkan secara lebih langsung dengan domain akuisisi, bukan dengan ketimpangan jumlah sampel. Jumlah 100 citra dipilih sebagai ukuran eksperimen awal yang memungkinkan variasi pencahayaan, bayangan, kontras, blur, sudut pandang, dan derau tetap terwakili, sekaligus mempertimbangkan beban penyusunan serta pemeriksaan ground truth tingkat piksel.

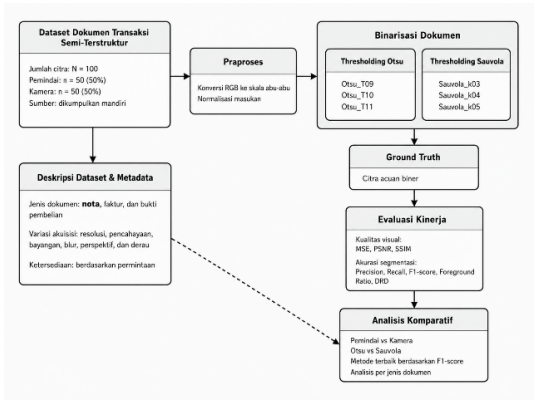
Setiap citra diuji menggunakan enam konfigurasi, sehingga eksperimen menghasilkan 600 keluaran binarisasi yang dibandingkan dengan 100 *ground truth*. Jumlah tersebut memadai untuk analisis deskriptif-komparatif dan penetapan baseline pada data uji ini, tetapi tidak ditetapkan melalui analisis daya statistik dan tidak dimaksudkan untuk mewakili seluruh variasi dokumen transaksi atau perangkat akuisisi.



Gambar 1. Contoh Dokumen Transaksi yang Diperoleh Menggunakan Kamera



Gambar 2. Contoh Dokumen Transaksi yang Diperoleh Menggunakan *Scanner*



Gambar 3. Gambaran Umum Metodologi Eksperimen

Tahap praproses dilakukan untuk menyiapkan citra sebelum binarisasi. Setiap citra berwarna terlebih dahulu dikonversi ke format skala abu-abu karena Otsu dan Sauvola bekerja berdasarkan intensitas piksel. Pada citra *scanner*, tahap ini berfungsi menstandarkan masukan yang relatif seragam. Pada citra kamera, tahap ini mempertahankan variasi intensitas yang timbul akibat bayangan, pencahayaan tidak merata, dan kontras lokal sehingga perbandingan metode tetap mencerminkan perbedaan kondisi akuisisi

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

ISSN: 2598-8565 (media cetak)
ISSN: 2620-4339 (media online)

Tabel 1. Parameter Perbandingan dan Standar Evaluasi

Kelompok Parameter	Metrik	Arah Keputusan	Fungsi dalam Pengolahan Dokumen Transaksi UMKM
Kualitas visual	MSE	Semakin rendah semakin baik	Mengukur kesalahan tingkat piksel terhadap ground truth.
Kualitas visual	PSNR	Semakin tinggi semakin baik	Menunjukkan tingkat kualitas citra dan perbedaan derau.
Kualitas visual	SSIM	Semakin tinggi semakin baik	Menunjukkan kemiripan struktural komponen dokumen.
Akurasi segmentasi	*Precision*	Semakin tinggi semakin baik	Mencegah bayangan, noda, atau area latar belakang terbaca sebagai informasi transaksi.
Akurasi segmentasi	*Recall*	Semakin tinggi semakin baik	Mempertahankan nama barang, nilai nominal, total, dan garis tabel.
Akurasi segmentasi	*F1-score*	Semakin tinggi semakin baik. Standar utama.	Menentukan metode paling seimbang untuk praproses sebelum OCR atau penguraian dokumen.
Pengendalian foreground	*Foreground Ratio*	Diharapkan bernilai seimbang	Menunjukkan apakah metode menghasilkan piksel foreground yang terlalu sedikit atau terlalu banyak.
Distorsi struktural	DRD	Semakin rendah semakin baik	Menilai distorsi visual yang dapat memengaruhi keterbacaan dokumen.

Evaluasi dilakukan pada dua tingkat. MSE, PSNR, dan SSIM digunakan untuk menjelaskan kedekatan visual terhadap *ground truth*, sedangkan *Precision, Recall, F1-score, Foreground Ratio*, dan DRD digunakan untuk menilai segmentasi *foreground* serta distorsi struktural. *F1-score* menjadi kriteria utama karena menyeimbangkan piksel *foreground* yang benar-benar dipertahankan dan derau latar belakang yang keliru tersegmentasi; DRD digunakan sebagai pendamping ketika dua hasil memiliki kualitas segmentasi yang berdekatan.

Istilah metode terbaik pada penelitian ini berarti konfigurasi dengan F1-score tertinggi di dalam domain akuisisi dan rentang parameter yang diuji, bukan optimum global Otsu atau Sauvola.Analisis kegagalan dilakukan secara deskriptif dengan menghubungkan pola *Precision, Recall, Foreground Ratio*, dan DRD dengan kerusakan visual berupa karakter terputus, hilangnya angka atau garis tipis, penebalan karakter, serta derau latar belakang. Analisis ini diprioritaskan pada citra kamera karena hasilnya paling rendah dan paling sensitif terhadap pencahayaan, bayangan, blur, serta distorsi perspektif.

Binarisasi Citra

Tahap binarisasi bertujuan mengubah citra hasil pemotongan menjadi citra dua tingkat, yaitu hitam dan putih, sehingga area teks dan komponen dokumen dapat tampak lebih jelas terhadap latar belakang. Penelitian ini menggunakan dua metode utama, yaitu *thresholding* Otsu dan *thresholding* Sauvola, yang masing-masing merepresentasikan pendekatan *thresholding* global dan *thresholding* lokal adaptif.

Metode Otsu

Metode Otsu memilih nilai ambang global yang memaksimalkan varians antar-kelas intensitas *foreground* dan latar belakang (Liu et al., 2024; Jumiawi & El-Zaart, 2022). *Varians* antar-kelas dinyatakan sebagai berikut:

$$\sigma^2_B = \omega_0\omega_1(\mu_0 - \mu_1)^2$$

dengan ω_0 dan ω_1 sebagai probabilitas kelas piksel latar belakang dan foreground, serta μ_0 dan μ_1 sebagai intensitas rata-rata masing-masing kelas. Nilai ambang Otsu dimodifikasi menggunakan faktor skala α:

$$T = \alpha \times T_Otsu, \quad \alpha \in \{0.9, 1.0, 1.1\}$$

Tiga konfigurasi representatif yang digunakan adalah Otsu_T09 (α = 0.9), Otsu_T10 (α = 1.0 sebagai ambang Otsu standar), dan Otsu_T11 (α = 1.1). Rentang ±10% dipilih untuk menguji perubahan terkendali di sekitar ambang otomatis: α = 0.9 merepresentasikan penurunan ambang, α = 1.0 menjadi baseline, dan α = 1.1 merepresentasikan peningkatan ambang. Tiga nilai tersebut juga menjaga jumlah konfigurasi Otsu sebanding dengan Sauvola. Pengujian ini merupakan analisis sensitivitas parameter terbatas, bukan optimasi global seluruh faktor skala.

Metode Sauvola

Thresholding Sauvola memperluas binarisasi lokal adaptif dengan menggunakan nilai rata-rata dan simpangan baku lokal pada lingkungan piksel (Naik, 2024; Li et al., 2021). Nilai ambang dihitung sebagai berikut:

$$T(x, y) = m(x, y)\ [1 + k\ (s(x, y)/R - 1)]$$

Pada persamaan tersebut, m(x,y) adalah rata-rata lokal, s(x,y) adalah simpangan baku lokal, R adalah rentang dinamis simpangan baku, dan k

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

ISSN: 2598-8565 (media cetak)
ISSN: 2620-4339 (media online)

mengatur sensitivitas ambang lokal. Nilai k = 0.3, 0.4, dan 0.5 dipilih sebagai tiga tingkat sensitivitas yang meningkat secara teratur. Ukuran jendela dipertahankan pada 25 × 25 piksel agar perubahan hasil terutama dapat dikaitkan dengan nilai k, bukan perubahan lingkungan lokal.

Peningkatan Kontras Menggunakan CLAHE

Contrast Limited Adaptive Histogram Equalization (CLAHE) diterapkan setelah konversi skala abu-abu dan sebelum binarisasi untuk meningkatkan kontras lokal sekaligus membatasi penguatan derau yang berlebihan (Kamel et al., 2024; Tung, 2021). Tahap ini sangat relevan bagi citra kamera yang dipengaruhi oleh bayangan, kontras rendah, dan variasi pencahayaan lokal.

Penyusunan *Ground Truth*

Sebanyak 100 *ground truth* biner dibuat secara manual, masing-masing satu untuk setiap citra, dengan resolusi dan dimensi yang sama seperti citra masukan. Teks, angka, simbol, informasi nominal, garis tabel, dan batas kolom diberi label *foreground*, sedangkan kertas, ruang kosong, noda, bayangan, serta area di luar dokumen diberi label *background*. Anotasi diperiksa pada pembesaran 200–400% untuk menjaga kontinuitas goresan tipis dan mencegah bayangan atau noda masuk sebagai *foreground.*

Validasi manual dilakukan dalam dua tahap oleh tim peneliti: pertama, *mask* ditumpangtindihkan dengan citra asli untuk memeriksa pergeseran, karakter terputus, digit yang hilang, dan kebocoran latar belakang; kedua, seluruh pasangan citra-*mask* diperiksa kembali setelah anotasi selesai untuk memastikan ukuran, orientasi, dan konvensi label konsisten. Kasus yang tidak konsisten diperbaiki sebelum penghitungan metrik. Karena validasi belum menggunakan penilai independen dan ukuran kesepakatan antar-anotator, kemungkinan subjektivitas ground truth tetap menjadi keterbatasan penelitian.

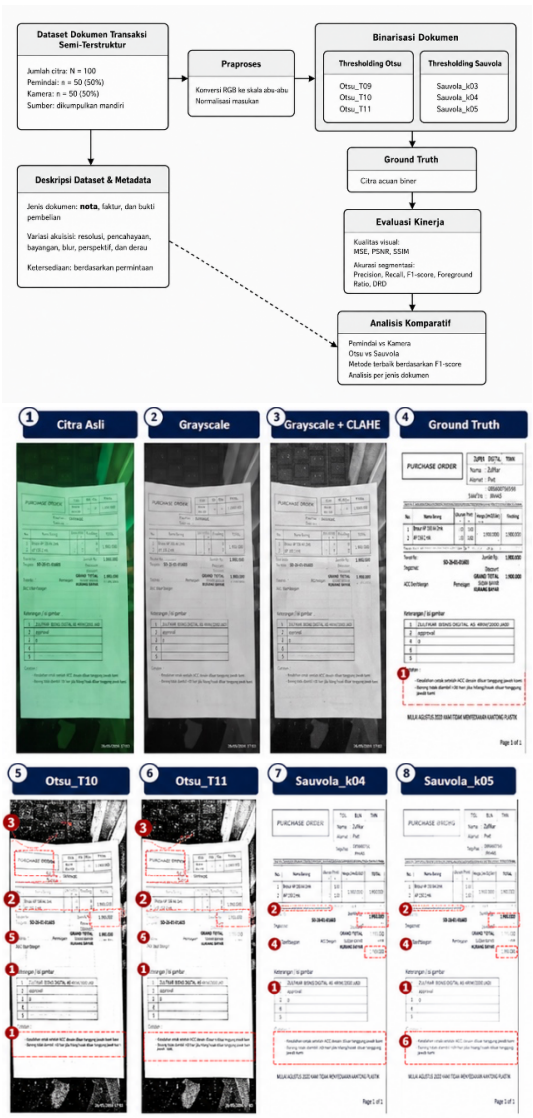
Komponen dan Analisis Dokumen Semi-Terstruktur

Dokumen semi-terstruktur dianalisis sebagai gabungan teks, angka, simbol, garis tabel, kolom, dan ruang kosong. Setiap citra diproses dengan Otsu_T09, Otsu_T10, Otsu_T11, Sauvola_k03, Sauvola_k04, dan Sauvola_k05 menggunakan prosedur yang sama pada kedua domain.

Hasil dinilai berdasarkan kemampuan mempertahankan karakter dan struktur penting, sekaligus menghindari *foreground* berlebihan, karakter menebal, serta derau latar belakang. Analisis dilakukan secara deskriptif-komparatif pada tingkat konfigurasi dan domain. Otsu diperlakukan sebagai *baseline* global, sedangkan Sauvola sebagai baseline lokal-adaptif. Interpretasi akhir menggabungkan peringkat metrik dengan pemeriksaan pola kerusakan visual pada hasil binarisasi.

HASIL DAN PEMBAHASAN

Pengujian dilakukan untuk mengetahui perbedaan hasil binarisasi antara thresholding Otsu dan thresholding Sauvola pada citra nota semi-terstruktur yang diperoleh melalui scanner kamera. Citra scanner merepresentasikan kondisi visual yang relatif terkendali, sedangkan citra kamera merepresentasikan kondisi dengan variasi pencahayaan, bayangan, derau kamera, sedikit blur, dan perubahan sudut pengambilan.



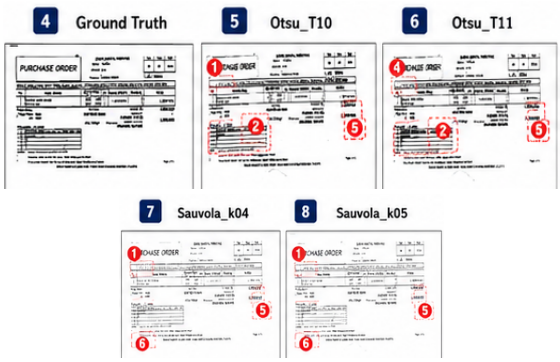
Gambar 4. Proses yang dilakukan Berdasrkan Metodologi Eksperimen

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

ISSN: 2598-8565 (media cetak)
ISSN: 2620-4339 (media online)



Gambar 5. Hasil Praproses dan Binarisasi Dokumen Kamera



Gambar 6. Hasil Praproses dan Binarisasi Dokumen pada Domain *Scanner*

Seluruh 100 citra diproses menggunakan enam konfigurasi yang sama dan dibandingkan dengan ground truth. Ringkasan metrik pada Tabel 2 dan Tabel 3 selanjutnya dibaca bersama contoh visual pada Gambar 5 dan Gambar 6 agar perubahan nilai metrik dapat dijelaskan melalui pola kerusakan karakter dan latar belakang.

Pada domain *scanner* (Gambar 6), Otsu_T09 menghasilkan latar relatif bersih, tetapi sebagian goresan tipis, ujung huruf, digit kecil, dan segmen garis tabel tampak terputus atau hilang; pola ini konsisten dengan Recall 0.747. Otsu_T10 mempertahankan kontinuitas karakter dan garis secara lebih seimbang. Ketika ambang dinaikkan menjadi Otsu_T11, lebih banyak goresan terdeteksi, tetapi karakter menjadi lebih tebal, jarak antargoresan menyempit, dan sebagian komponen berdekatan cenderung menyatu, sehingga Precision turun menjadi 0.750.

Sauvola, kenaikan k dari 0.3 ke 0.5 membuat hasil semakin selektif: latar belakang bersih dan Precision mencapai 1.000, tetapi goresan tipis, bagian angka, serta garis tabel semakin terfragmentasi. Penurunan Recall dari 0.584 menjadi 0.322 menunjukkan bahwa kebersihan visual tersebut diperoleh dengan kehilangan foreground yang penting. Pada domain kamera (Gambar 3), bayangan, tepi dokumen yang gelap, dan area di luar bidang kertas membentuk blok foreground tebal pada hasil Otsu.

Peningkatan ambang dari Otsu_T09 ke Otsu_T11 memperluas area hitam tersebut, tercermin pada Foreground Ratio yang meningkat dari 0.339 menjadi 0.375 dan Precision yang tetap sangat rendah (0.121–0.124). Sauvola lebih mampu menekan blok latar belakang; Sauvola_k03 bahkan menghasilkan DRD terendah. Namun, ambang lokal menghapus goresan berkontras rendah sehingga huruf dan angka tampak terputus, titik atau tanda baca hilang, dan garis tabel tidak berkesinambungan. Ketika k meningkat, Recall turun dari 0.082 menjadi 0.052. Dengan demikian, citra yang tampak lebih bersih belum tentu lebih akurat secara segmentasi, karena elemen transaksi yang seharusnya dipertahankan justru hilang.

Tabel 2. Evaluasi Visual Otsu dan Sauvola pada Domain *Scanner* dan Kamera.

Domain Akuisisi	Konfigurasi Metode	MSE (semakin rendah semakin baik)	PSNR dalam dB (semakin tinggi semakin baik)	SSIM (semakin tinggi semakin baik)
Scanner	Otsu_T09	671.466	19.861	0.921
Scanner	Otsu_T10	945.617	18.374	0.916
Scanner	Otsu_T11	1399.767	16.670	0.905
Scanner	Sauvola_k03	544.393	20.772	0.921
Scanner	Sauvola_k04	463.416	21.471	0.917
Scanner	Sauvola_k05	433.344	21.762	0.911
Kamera	Otsu_T09	4327.527	11.768	0.511
Kamera	Otsu_T10	4256.142	11.841	0.513
Kamera	Otsu_T11	4296.650	11.800	0.514
Kamera	Sauvola_k03	14426.866	6.539	0.601
Kamera	Sauvola_k04	14735.990	6.447	0.599
Kamera	Sauvola_k05	14954.769	6.383	0.593

Tabel 3. Evaluasi Segmentasi Otsu dan Sauvola pada Domain *Scanner* dan Kamera

Domain Akuisisi	Konfigurasi Metode	Precision (semakin tinggi semakin baik)	Recall (semakin tinggi semakin baik)	F1-score (standar utama)	Foreground Ratio	DRD (semakin rendah semakin baik)
Scanner	Otsu_T09	0.991	0.747	0.852	0.026	0.260
Scanner	Otsu_T10	0.914	0.921	0.917	0.035	0.166
Scanner	Otsu_T11	0.750	0.994	0.855	0.046	0.337
Scanner	Sauvola_k03	1.000	0.584	0.737	0.020	0.416

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

ISSN: 2598-8565 (media cetak)
ISSN: 2620-4339 (media online)

Domain Akuisisi	Konfigurasi Metode	Precision (semakin tinggi semakin baik)	Recall (semakin tinggi semakin baik)	F1-score (standar utama)	Foreground Ratio	DRD (semakin rendah semakin baik)
Scanner	Sauvola_k04	1.000	0.432	0.604	0.015	0.568
Scanner	Sauvola_k05	1.000	0.322	0.487	0.011	0.678
Kamera	Otsu_T09	0.121	0.246	0.162	0.339	2.544
Kamera	Otsu_T10	0.122	0.263	0.167	0.358	2.625
Kamera	Otsu_T11	0.124	0.280	0.172	0.375	2.693
Kamera	Sauvola_k03	0.155	0.082	0.107	0.088	0.136
Kamera	Sauvola_k04	0.154	0.064	0.091	0.070	1.289
Kamera	Sauvola_k05	0.154	0.052	0.078	0.057	1.236

Pada *scanner*, Sauvola_k05 menghasilkan MSE terendah (433.3445) dan PSNR tertinggi (21.7625), sedangkan Sauvola_k03 mencapai SSIM 0.9210. Akan tetapi, keunggulan visual tersebut disertai kehilangan foreground. Otsu_T10 memberikan kompromi yang lebih fungsional antara kualitas visual dan keterjagaan karakter.

Pada kamera, Otsu_T10 menghasilkan MSE terendah (4256.1416) dan PSNR tertinggi (11.8406), sementara Sauvola_k03 mencapai SSIM tertinggi (0.6009). Perbedaan arah metrik ini menegaskan bahwa kemiripan struktur atau kebersihan visual tidak dapat digunakan sendiri untuk menentukan kesiapan hasil bagi OCR.

Pada *scanner*, Otsu_T10 menjadi konfigurasi paling seimbang dengan Precision 0.9136, Recall 0.9208, F1-score 0.9172, Foreground Ratio 0.0349, dan DRD 0.1663. Otsu_T09 lebih presisi tetapi kehilangan lebih banyak goresan, sedangkan Otsu_T11 mempertahankan hampir seluruh foreground tetapi meningkatkan penebalan dan false positive.

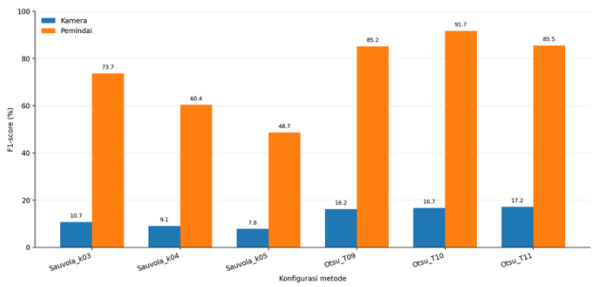
$$F1 = 2 \times (0.9136 \times 0.9208) / (0.9136 + 0.9208) = 0.9172$$

Pada kamera, seluruh konfigurasi gagal mencapai segmentasi yang andal. Otsu_T11 memperoleh F1-score tertinggi (0.1720), tetapi Precision 0.1242 dan Recall 0.2798 menunjukkan banyak derau sekaligus banyak foreground yang hilang. Sauvola_k03 memiliki DRD 0.1364 dan SSIM 0.6009, namun F1-score hanya 0.1070 karena Recall sangat rendah.

Tren Sauvola konsisten pada kedua domain: peningkatan k membersihkan latar dan mempertahankan Precision, tetapi mengurangi Recall karena semakin banyak goresan tipis dan struktur tabel terhapus. Oleh sebab itu, Sauvola_k03 adalah konfigurasi Sauvola terbaik dalam rentang yang diuji, bukan parameter optimal secara umum

Tabel 4. Perbandingan *F1-score* terbaik berdasarkan domain akuisisi

Domain Akuisisi	Otsu Terbaik Berdasarkan F1-score	Sauvola Terbaik Berdasarkan F1-score	Selisih F1-score	Fungsi yang Direkomendasikan dalam Alur Kerja Transaksi UMKM
Scanner	Otsu_T10 = 0.917	Sauvola_k03 = 0.737	0.180	Otsu_T10 sebagai metode praproses utama untuk dokumen transaksi berbasis *scanner* karena memberikan segmentasi *foreground* yang seimbang.
Kamera	Otsu_T11 = 0.172	Sauvola_k03 = 0.107	0.065	Otsu_T11 hanya sebagai opsi dengan *F1-score* lebih baik, tetapi tambahkan praproses pencahayaan dan perspektif. Gunakan Sauvola_k03 apabila distorsi visual yang lebih rendah diprioritaskan.
Rata-rata	0.545	0.422	0.123	Otsu lebih sesuai sebagai metode utama yang berorientasi pada segmentasi dalam data uji ini.



Gambar 6. Perbandingan *F1-score* Berdasarkan Domain Akuisisi dan Konfigurasi Metode

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Tabel 4 menunjukkan bahwa Otsu menghasilkan *F1-score* tertinggi pada kedua domain akuisisi. Selisih antara Otsu terbaik dan Sauvola terbaik adalah 0.1802 pada domain *scanner* dan 0.0650 pada domain kamera. Perbedaan tersebut menunjukkan bahwa pada data uji yang digunakan, pendekatan *thresholding* global Otsu lebih efektif dalam mempertahankan keseimbangan *Precision* dan *Recall*. Namun, rendahnya *F1-score* pada domain kamera tetap menunjukkan bahwa akuisisi kamera memerlukan penanganan tambahan sebelum binarisasi.

Perbedaan kinerja berkaitan dengan karakteristik mendasar kedua metode. Otsu menggunakan satu nilai ambang global sehingga lebih stabil ketika distribusi intensitas *foreground* dan latar belakang terpisah secara memadai. Sauvola menghitung nilai ambang lokal agar lebih responsif terhadap variasi intensitas, tetapi parameter lokal yang terlalu selektif dapat menurunkan *Recall*. Pada dokumen faktur, penurunan *Recall* berisiko menghilangkan karakter tipis, angka kecil, atau garis tabel penting pada tahap pengenalan teks.

Secara keseluruhan, Otsu_T10 paling sesuai untuk *scanner*, sedangkan Otsu_T11 hanya menjadi konfigurasi baseline terbaik pada kamera. Nilai kamera yang tetap rendah menunjukkan bahwa persoalan utama tidak dapat diselesaikan hanya dengan mengubah ambang; distorsi geometris dan fotometris harus ditangani sebelum binarisasi.

Hasil penelitian ini berbeda dari Darpito et al. (2026), yang mengevaluasi 200 dokumen terstruktur SATS-LN (150 *scanner* dan 50 kamera), menggunakan pelokalan tujuh bidang dengan Detectron2, dan melaporkan kinerja kamera terbaik pada Sauvola_k04 dengan F1-score 0.980. Pada penelitian ini, F1-score kamera tertinggi hanya 0.172 karena dokumen transaksi bersifat semi-terstruktur, tata letaknya lebih bervariasi, dan metode diuji sebagai baseline tanpa pelokalan bidang serta koreksi perspektif yang kuat. Perbedaan tersebut tidak menunjukkan pertentangan antarmetode, melainkan memperlihatkan bahwa pemotongan wilayah dokumen dan normalisasi akuisisi sangat menentukan hasil thresholding.

Kajian Yang et al. (2024) menempatkan variasi pencahayaan, bercak, patahan goresan, kemiringan, dan deformasi sebagai tantangan utama binarisasi dokumen. Sejalan dengan kajian tersebut, kontribusi penelitian ini bukan algoritma binarisasi baru, tetapi baseline lintas-domain yang seimbang pada dokumen transaksi UMKM, penggunaan metrik visual dan segmentasi secara bersamaan, serta analisis kegagalan yang menghubungkan karakter terputus dan derau tebal dengan kesiapan OCR. Kebaruan praktisnya adalah menunjukkan secara empiris batas penggunaan Otsu dan Sauvola dasar pada citra kamera serta menetapkan pemotongan/koreksi perspektif, normalisasi iluminasi, dan reduksi bayangan sebagai tahap wajib dalam pipeline lanjutan.

Implikasi bagi Pengolahan Transaksi UMKM

Untuk akuisisi *scanner*, Otsu_T10 dapat digunakan sebagai baseline praproses karena menjaga karakter, angka, dan garis tabel secara relatif seimbang. Rekomendasi ini terbatas pada kondisi *scanner* dan konfigurasi yang diuji. Untuk akuisisi kamera, thresholding tidak boleh diterapkan langsung pada citra mentah. Pemotongan atau koreksi perspektif, normalisasi iluminasi, reduksi bayangan, dan penyelarasan citra terhadap ground truth perlu dilakukan lebih dahulu agar tepi gelap dan bayangan tidak menjadi foreground, serta goresan berkontras rendah tidak hilang. Dengan demikian, keluaran penelitian ini digunakan sebagai baseline untuk merancang pipeline kamera yang lebih lengkap, bukan sebagai bukti bahwa thresholding dasar sudah memadai bagi OCR langsung.

KESIMPULAN

Penelitian ini menunjukkan bahwa kinerja Otsu dan Sauvola sangat bergantung pada domain akuisisi dan konfigurasi parameter. Pada *scanner*, Otsu_T10 memberikan hasil paling seimbang dengan Precision 0.9136, Recall 0.9208, F1-score 0.9172, dan DRD 0.1663. Pada kamera, Otsu_T11 memperoleh F1-score tertinggi sebesar 0.1720, sedangkan Sauvola_k03 menghasilkan DRD terendah sebesar 0.1364 tetapi F1-score hanya 0.1070. Analisis visual memperlihatkan bahwa *scanner* terutama mengalami karakter tipis yang terputus atau karakter yang menebal, sementara kamera mengalami blok derau tebal, karakter terfragmentasi, digit kecil yang hilang, dan garis tabel yang terputus.

Otsu_T10 direkomendasikan hanya sebagai baseline untuk citra *scanner* pada rentang parameter yang diuji. Otsu_T11 dan Sauvola_k03 belum layak digunakan langsung untuk OCR kamera. Pada citra kamera, pemotongan atau koreksi perspektif, normalisasi iluminasi, dan reduksi bayangan merupakan tahap wajib sebelum thresholding. Keterbatasan penelitian meliputi 100 citra yang dikumpulkan mandiri, rentang parameter yang terbatas, satu ukuran jendela Sauvola, validasi ground truth tanpa ukuran kesepakatan antar-anotator, dan belum adanya evaluasi OCR hilir. Penelitian berikutnya perlu memperluas variasi perangkat serta dokumen, menerapkan validasi ground truth independen, menguji kombinasi parameter yang lebih luas, dan

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

membandingkan pipeline baseline dengan preprocessing geometris-fotometris serta metrik CER/WER.

DAFTAR PUSTAKA

Abbas, S. A., & Hussain, S. ul. (2017). Recovering homography from camera captured documents using convolutional neural networks. *arXiv*. https://arxiv.org/abs/1709.03524

Bella, F., El Rhabi, M., Hakim, A., & Laghrib, A. (2022). An innovative document image binarization approach driven by the non-local p-Laplacian. *EURASIP Journal on Advances in Signal Processing, 2022*. https://doi.org/10.1186/s13634-022-00883-2

Bloechle, J.-L., Hennebert, J., & Gisler, C. (2023). YinYang, a fast and robust adaptive document image binarization for optical character recognition. *Proceedings of the ACM Symposium on Document Engineering 2023*. https://doi.org/10.1145/3573128.3609354

Darpito, M. N., Firdausy, K., & Fadlil, A. (2026). Performance evaluation of Otsu and Sauvola thresholding for structured document binarization. *Scientific Journal of Informatics, 13*(1), 189–200. https://doi.org/10.15294/sji.v13i1.40245

Das, S., Ma, K., Shu, Z., Samaras, D., & Shilkrot, R. (2019). DewarpNet: Single-image document unwarping with stacked 3D and 2D regression networks. P*roceedings of the IEEE/CVF International Conference on Computer Vision*, 131–140. https://doi.org/10.1109/ICCV.2019.00022

Feng, H., Wang, Y., Zhou, W., Deng, J., & Oct, C. V. (2021). DocTr: Document image transformer for geometric unwarping and illumination correction. *Association for Computing Machinery.* https://doi.org/10.1145/3474085.3475388

Gatos, B., Ntirogiannis, K., & Pratikakis, I. (2009). ICDAR 2009 document image binarization contest (DIBCO 2009). *2009 10th International Conference on Document Analysis and Recognition,* 1375–1382. https://doi.org/10.1109/ICDAR.2009.246

Gatos, B., Pratikakis, I., & Perantonis, S. J. (2006). Adaptive degraded document image binarization. *Pattern Recognition, 39*(3), 317–327. https://doi.org/10.1016/j.patcog.2005.09.010

Guan, S., et al. (2025). PreP-OCR: A complete pipeline for document image restoration and enhanced OCR accuracy, 1, 15413–15425.

Ha, H. T., & Horak, A. (2022). Information extraction from scanned invoice images using text analysis and layout features. *Signal Processing: Image Communication, 102*, 116601. https://doi.org/10.1016/j.image.2021.116601

Han, S., Ji, S., & Rhee, J. (2023). Diffusion-denoising process with gated U-Net for high-quality document binarization. *Applied Sciences*. https://doi.org/10.20944/preprints202308.2048.v1

Huang, Z., Chen, K., He, J., Bai, X., Karatzas, D., Lu, S., & Jawahar, C. V. (2019). ICDAR2019 competition on scanned receipt OCR and information extraction. *2019 International Conference on Document Analysis and Recognition,* 1516–1520. https://doi.org/10.1109/ICDAR.2019.00244

Jing, Z., & Tang, B. (2024). Improved image segmentation method based on Otsu thresholding and level set techniques. *Journal of Physics: Conference Series, 2813*, 012017. https://doi.org/10.1088/1742-6596/2813/1/012017

Jumiawi, W. A. H., & El-Zaart, A. (2022). Improvement in the between-class variance based on lognormal distribution for accurate image segmentation.

Kamel, O., Amin, K. M., Semary, N. A., & Aboelenien, N. (2024). An automated contrast enhancement technique for remote sensed images. IJCI. *International Journal of Computers and Information, 11*(1), 1-16.

Kasem, M. S., Mahmoud, M., Senussi, M. F., Abdalla, M., & Kang, H. S. (2026). KORIE: A multi-task benchmark for detection, OCR, and information extraction on Korean retail receipts. *Mathematics, 14*(1). https://doi.org/10.3390/math14010187

Li, D., Wu, Y., & Zhou, Y. (2021). SauvolaNet: Learning adaptive Sauvola network for degraded document binarization. *arXiv*. https://doi.org/10.1007/978-3-030-86337-1_36

Li, H., & Yu, X. (2025). Intelligent management accounting tools based on big data algorithms and invoice automatic recognition algorithms. *Journal of Computational Methods in Sciences and Engineering, 25*(5), 4335–4347. https://doi.org/10.1177/14727978251337886

Liu, X., Zhang, Z., Hao, Y., Zhao, H., & Yang, Y. (2024). Temperature feature extraction method for infrared images.

Naik, Y. (2024). Enhancement of degraded historical document images for binarization. *Journal of Electrical Systems.* https://doi.org/10.52783/jes.5990

Nikolaidou, K., Seuret, M., Mokayed, H., & Liwicki, M. (2022). A survey of historical document image datasets. *International Journal on Document Analysis and Recognition, 25*(4), 305–338. https://doi.org/10.1007/s10032-022-00405-8

Pratikakis, I., Zagoris, K., Barlas, G., & Gatos, B. (2017). ICDAR2017 competition on document image binarization (DIBCO 2017). *2017 14th IAPR International Conference on Document*

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

ISSN: 2598-8565 (media cetak)
ISSN: 2620-4339 (media online)

Analysis and Recognition, 1395–1403. https://doi.org/10.1109/ICDAR.2017.228

Quattrini, F., Pippi, V., Cascianelli, S., & Cucchiara, R. (2024). Binarizing documents by leveraging both space and frequency. *arXiv*. https://doi.org/10.48550/arXiv.2404.17243

Thendral, R., & Abitha, S. (2026). AutoBill voice assistant: Intelligent invoice reader automates data extraction and provides real-time audio feedback. *2026 IEEE International Students' Conference on Electrical, Electronics and Computer Science*. https://doi.org/10.1109/SCEECS68810.2026.11429800

Tung, T. (2021). ICEBIN: Image contrast enhancement based on induced norm and local patch approaches. *IEEE Access, 9*, 23737–23750. https://doi.org/10.1109/ACCESS.2021.3056244

Xiong, W., Zhou, L., Yue, L., Li, L., & Wang, S. (2021). An enhanced binarization framework for degraded historical document images. *EURASIP Journal on Image and Video Processing*, 2021. https://doi.org/10.1186/s13640-021-00556-4

Yang, Z., Zuo, S., Zhou, Y., He, J., & Shi, J. (2024). A review of document binarization: Main techniques, new challenges, and trends. *Electronics, 13*(7), 1394. https://doi.org/10.3390/electronics13071394

Zhang, L., Wang, K., & Wan, Y. (2024). An efficient transformer–CNN network for document image binarization. *Electronics, 13*(12), 2243. https://doi.org/10.3390/electronics13122243

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.