

Identifikasi Text Meteran Air Menggunakan Metode Run-Length Smearing Algorithm (Rlsa)

Marissa Utami ^(✉)

Universitas Muhammadiyah Bengkulu, Bengkulu, Indonesia
marissautami@umb.ac.id

Erwin Dwika Putra¹, Marissa Utami², Muhammad Husni Rifqo³

Universitas Muhammadiyah Bengkulu, Bengkulu, Indonesia
erwindwikap@gmail.com

Abstrak— Proses pencatatan yang dilakukan oleh petugas perusahaan air minum masih manual yaitu menggunakan pencatatan dengan menggunakan sebuah buku, sehingga mengakibatkan proses pencatatan meter membutuhkan waktu yang lama. Sistem pembacaan manual masih diterapkan dalam proses pembacaan pemakaian air pelanggan. Hal ini menyebabkan proses yang terjadi memakan banyak waktu. Keluhan-keluhan konsumen pun meningkat dengan ditemukannya beberapa petu-gas pencatat yang melakukan kesalahan dalam proses pencatatan. Solusi yang dapat dilakukan untuk dapat mengatasi masalah pencatatan angka pemakaian pelanggan adalah dengan membuat aplikasi untuk membaca angka pemakaian pelanggan pada meteran air. Salah satu cara yang dapat dilakukan un-tuk mengenali karakter adalah dengan Optical Character Recognition (OCR). Identifikasi Angka Meter Air Dengan Metode Run-Length Smearing Algorithm ber-hasil penulis bangun dan Metode RLSA dapat mengenali pola karakter angka pada citra Angka Meter Air, dan berdasarkan tingkat pengukuran dari hasil penelitian yang dilakukan yaitu menunjukkan tingkat keberhasilan identifikasi Anka meteran air yaitu 83%.

Abstract— *The recording process carried out by drinking water company officers is still manual, namely using records using a book, resulting in the meter recording process taking a long time. The manual reading system is still applied in the process of reading customer water usage. This causes the process to take a lot of time. Consumer complaints also increased with the discovery of several recording officers who made mistakes in the recording process. The solution that can be done to overcome the problem of recording customer usage figures is to create an application to read customer usage figures on the water meter. One way that can be done to recognize characters is with Optical Character Recognition (OCR). The identification of the Water Meter Number Using the Run-Length Smearing Algorithm method was successful, the authors built and the RLSA Method was able to recognize the pattern of numeric characters in the Water Meter Number image, and based on the level of measurement from the results of the research carried out, it showed the success rate of identification. The water meter rate is 83%.*

Keywords: ocr, rlsa, meter, water.

1 Pendahuluan

Alat untuk mengukur banyaknya aliran air secara terus menerus melalui sistem kerja peralatan yang dilengkapi dengan unit sensor, unit penghitung, dan unit indikator pengukur untuk menyatakan volume air yang lewat. Sedangkan badan meteran air merupakan bagian utama yang merupakan ruang untuk menempatkan alat hitung dan mempunyai saluran masuk dan saluran keluar pada sisi yang berlawanan. (SNI 2547:2008).

Proses pencatatan yang dilakukan oleh petugas perusahaan air minum masih manual yaitu menggunakan pencatatan dengan menggunakan sebuah buku, sehingga mengakibatkan proses pencatatan meter membutuhkan waktu yang lama. Sistem pembacaan manual masih diterapkan dalam proses pembacaan pemakaian air pelanggan. Hal ini menyebabkan proses yang terjadi memakan banyak waktu. Keluhan-keluhan konsumen pun meningkat dengan ditemukannya beberapa petugas pencatat yang melakukan kesalahan dalam proses pencatatan [1]. Dengan penelitian ini dikembangkan sebuah perangkat lunak yang bertujuan untuk memperbaiki metode pencatatan manual pada sistem, perbaikan yang diharapkan adalah mengurangi waktu proses pencatatan dan kesalahan pembacaan meteran air [2][3].

Optical Character Recognition adalah sebuah proses untuk mengkonversi dokumen cetak atau tulisan tangan hasil scan menjadi karakter ASCII (karakter yang dapat dibaca oleh mesin)[2][3][4]. Dengan kata lain, pengenalan teks secara otomatis menggunakan OCR adalah proses untuk mengkonversi citra dokumen teks menjadi teks berbentuk data digital sehingga dapat diedit. Alat yang dilengkapi OCR akan mempercepat proses input data, serta mengurangi kemungkinan kesalahan dalam proses Entry data.

OCR pada dasarnya ialah pengenalan karakter alphanumeric dari karakter tulisan tangan atau file maupun citra menjadi teks yang dapat diedit dan penerapan metode Run Length smearing Algorithm digunakan untuk mendeteksi area meter air dari citra hasil pemotretan meteran air [5][6].

Solusi yang dapat dilakukan untuk dapat mengatasi masalah pencatatan angka pemakaian pelanggan adalah dengan membuat aplikasi untuk membaca angka pemakaian pelanggan pada meteran air. Salah satu cara yang dapat dilakukan untuk mengenali karakter adalah dengan Optical Character Recognition (OCR).

2 Studi Literatur

2.1 Optical Character Recognition (OCR)

Optical Character Recognition adalah sebuah proses untuk mengkonversi dokumen cetak atau tulisan tangan hasil scan menjadi karakter ASCII (karakter yang dapat dibaca oleh mesin). Dengan kata lain, pengenalan teks secara otomatis menggunakan OCR adalah proses untuk mengkonversi citra dokumen teks menjadi teks berbentuk data digital sehingga dapat diedit. Komputer yang dilengkapi dengan OCR akan mempercepat proses memasukkan data, serta mengurangi kemungkinan kesalahan dalam proses memasukkan data. [2][7].

Optical character recognition (OCR) adalah sebuah sistem komputer yang dapat membaca huruf, baik yang berasal dari sebuah pencetak (printer atau mesin ketik) maupun yang berasal dari tulisan tangan. OCR adalah aplikasi yang menerjemahkan gambar karakter (image character) menjadi bentuk teks dengan cara menyesuaikan pola karakter per-baris dengan pola yang telah tersimpan dalam database aplikasi [5][8]. Hasil dari proses OCR adalah berupa teks sesuai dengan gambar output scanner dimana tingkat keakuratan penerjemahan karakter tergantung dari tingkat kejelasan gambar dan metode yang digunakan [9]. Secara umum blok diagram kerja OCR dapat dilihat pada gambar.

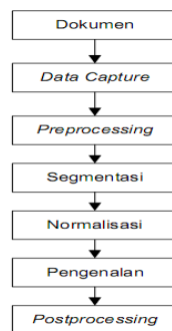


Fig.1. Struktur Optical Character Recognition (OCR)

1. Data Capture
Data capture merupakan proses konversi suatu dokumen (hardcopy) menjadi suatu file gambar digital.
2. Preprocessing
Preprocessing merupakan suatu proses untuk menghilangkan bagian-bagian yang tidak diperlukan pada gambar input untuk proses selanjutnya.
3. Segmentasi
Segmentasi adalah proses memisahkan area pengamatan (region) pada tiap karakter yang akan dideteksi.
4. Normalisasi
Normalisasi adalah proses merubah dimensi region tiap karakter. Dalam OCR algoritma yang digunakan pada proses ini adalah algoritma scaling.
5. Pengenalan
Pengenalan merupakan proses untuk mengenali karakter yang diamati dengan cara membandingkan ciri-ciri karakter yang diperoleh dengan ciri-ciri karakter yang ada pada database.
6. Postprocessing
Pada umumnya proses yang dilakukan pada tahap ini adalah proses koreksi ejaan sesuai dengan bahasa yang digunakan.

2.2 Run-Length Smearing Algorithm (RLSA)

Run-Length Smearing Algorithm (RLSA) atau biasa disebut dengan smearing merupakan salah satu metode yang berfungsi mencari lokasi teks dalam suatu gambar. Cara kerja metode ini adalah dengan cara melakukan proses scan-lines pada gambar

secara vertikal dan horisontal. Proses smearing ini dilakukan sebanyak tiga kali untuk mendapatkan lokasi teks yang benar. Pada smearing pertama kali dilakukan pencarian piksel hitam yang merupakan jarak dari tiap karakter huruf/angka, smearing kedua dilakukan pencarian piksel putih yang merupakan lebar huruf dari karakter. Setelah kedua smearing diatas dilakukan hasil dari keduanya kemudian digabungkan setelah itu dilakukan lagi smearing untuk menemukan piksel hitam yang merupakan besar dari lokasi teks tersebut secara utuh. Setelah lokasi angka didapatkan gambar tersebut akan diekstrak sehingga terbentuk suatu gambar yang berisikan angka saja [10][11]

Run-length Smearing Algorithm (RLSA) pada dasarnya di terapkan untuk binary sequence di mana pixel putih diwakilkan oleh 0 dan pixel hitam oleh 1. RLSA mengubah binary sequence x ke output sequence y berdasarkan aturan-aturan berikut :

1. jika jumlah 0 berurutan kurang dari atau sama dengan nilai C yang telah di definisikan 0 yang ada di x akan di ubah menjadi 1 di y .
2. 1 yang ada di x tidak berubah di y .

Sebagai contoh, dengan $C = 4$ sequence x akan di-mapping menjadi y sebagai berikut :

x : 00010000010100001000000011000

y : 1111000001111111000000011111

Jika diterapkan untuk array yang memiliki pattern, RLSA akan menghubungkan pixel-pixel hitam yang berdekatan yang terpisahkan oleh kurang dari C pixel. Dengan pemilihan nilai C yang tepat, pixel-pixel yang terhubung akan menjadi suatu area dengan tipe data tertentu. Pemilihan nilai C ini sangat penting. Nilai C yang terlalu kecil hanya akan menghubungkan pixel-pixel antar karakter itu sendiri. Nilai C yang agak lebih besar akan menggabungkan karakter-karakter menjadi kata, tetapi tidak cukup besar untuk menjembatani jarak antar dua kata. Secara umum, nilai C biasanya didapatkan dari tinggi karakter dan jarak antar baris. Ada dua macam cara untuk menjalankan metode RLSA yaitu, dijalankan secara horizontal atau dijalankan secara vertikal.

Setelah mendapatkan hasil RLSA, semua pixel hitam yang ada di posisi paling bawah akan diberi tanda. Pixel-pixel hitam itu akan memenuhi persamaan sudut antara dua titik seperti pada Rumus berikut :

$$a_i = \arctan \left(\frac{y_{i+1} - y_i}{x_{i+1} - x_i} \right)$$

$$\theta = \frac{\sum_{i=1}^{N-1} a_i}{N}$$

a_i = sudut antara 2 titik

y_i = posisi vertikal titik ke- i

x_i = posisi horizontal titik ke- i

θ = sudut Putar yang dicari

N = jumlah titik terbawah

Dengan menggunakan rumus $a_i = \arctan\left(\frac{y_{i+1}-y_i}{x_{i+1}-x_i}\right)$ akan didapatkan semua kemungkinan nilai sudut kemiringan gambar. Dari semua kemungkinan nilai sudut kemiringan gambar yang ada akan diambil nilai rata-rata sudut tersebut dengan menggunakan rumus $\theta = \frac{\sum_{i=1}^{N-1} a_i}{N}$.

3 Metodologi

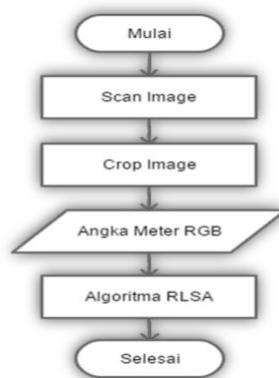


Fig.2. Alur Penelitian

Berdasarkan gambar diatas sistem ini terdapat beberapa tahapan yaitu mulai ketika menjalan sistem, scan image pada saat kamera sistem siap mengambil gambar setelah kamera mendapat hasil image yang di scan dilakukan cropping guna untuk memotong gambar yang tidak perlu, Angka meter RGB yang masih berwarna dilakukan proses Grayscale untuk mengubah warna RGB menjadi abu-abu putih, proses algoritma RLSA digunakan untuk mendeteksi lokasi angka dan mendapatkan hasil keluaran text. Metode pengujian yang digunakan pada penelitian ini adalah Mean Squared Error (MSE). Dalam statistik, Mean Squared Error (MSE) sebuah estimator adalah nilai yang diharapkan dari kuadrat error. Error yang ada menunjukkan seberapa besar perbedaan hasil estimasi dengan nilai yang akan diestimasi. Perbedaan itu terjadi karena adanya keacakan pada data atau karena estimator tidak mengandung informasi yang dapat menghasilkan estimasi yang lebih akurat.

$$MSE = \frac{1}{N} \sum_{t=h}^N (y_t - \hat{y}_t)^2$$

Di mana :

MSE = Mean Squared Error

N = Jumlah Sampel

y_t = Nilai Aktual Indeks

\hat{y}_t = Nilai Prediksi Indeks

4 Hasil dan Pembahasan

Hasil dari penelitian ini akan ditampilkan hasil capture dan Cropping, dimana tampilan dapat dilihat pada gambar 4.1, adapun fungsi dari cropping gambar di atas ialah untuk membantu sistem agar hasil deteksi lebih akurat.



Fig.3. Capture and Cropping

Pada gambar 4.2 merupakan tampilan dari menu Hasil Scan OCR, tampilan tersebut merupakan hasil dari deteksi angka yang telah dilakukan pada menu sebelumnya. Menu ini akan tampil secara otomatis ketika pengguna menekan tombol crop setelah pengguna selesai melakukan crop pada gambar yang akan di deteksi.

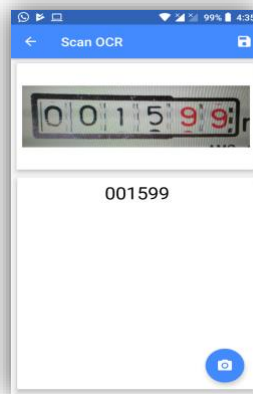


Fig.4. Hasil OCR RLSA

Pengujian

Untuk dapat melakukan pengujian MSE langkah pertama adalah membuka folder data gambar tersebut lalu tekan shift dan klik kanan pada mouse secara bersamaan lalu akan muncul konteks menu windows explorer dan pilih Open PowerShell Windows Home.

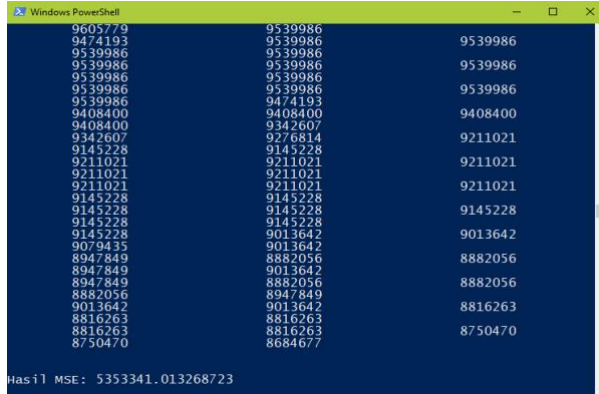


Fig.5. Hasil Pengujian MSE

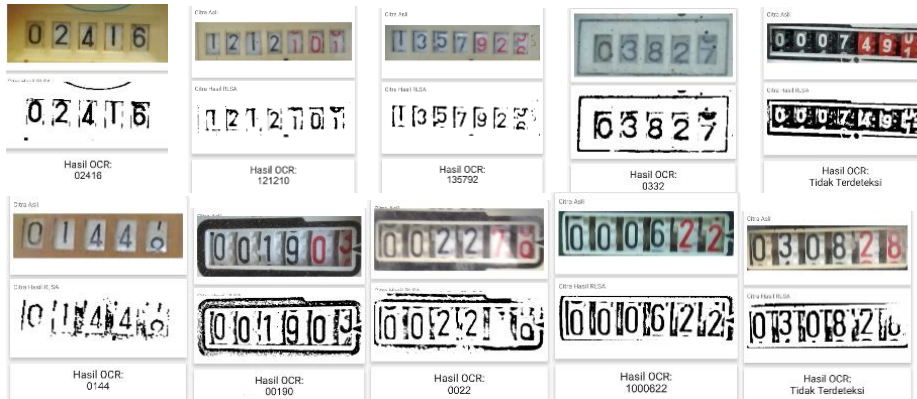


Fig.6. Hasil OCR RLSA

Total data yang digunakan pada penelitian ini adalah 90 data sampel gambar yang diambil langsung. Yang selanjutnya akan diukur tingkat akurasi keberhasilan penerapan algoritma RLSA OCR pada recognition text.

Data meter air tersebut akan dibandingkan dengan hasil klasifikasi yang dilakukan oleh sistem. Hasil pengujian akurasi klasifikasi data meter air dapat dilihat pada Tabel di bawah ini :

Table.1. Confusion Matrix

		Predicted Class	
		Positif	Negatif
Actual Class	Positif	70	10
	Negatif	5	5

Setelah sistem melakukan klasifikasi, lalu hitung nilai precision, recall dan akurasinya.

$$\text{Akurasi } \frac{P+N}{P+PN+NP+N} \times 100\% \quad \text{Precision } \frac{P}{P+PN} \times 100\% \quad \text{Recall } \frac{N}{N+NP} \times 100\%$$

$$\text{Akurasi} = \frac{70+5}{70+10+5+5} = 0,83$$

$$\text{Precision Positif} = \frac{70}{70+10} = 0,875$$

$$\text{Precision Negatif} = \frac{5}{5+5} = 0,5$$

Data pengujian akurasi yang digunakan sebanyak 90 data, yang terdiri dari 70 data positif dan 10 data negatif. Hasil klasifikasi yang dilakukan oleh sistem, sebanyak 75 data termasuk sentimen positif dan 15 termasuk sentimen negatif. Berdasarkan pengujian akurasi, didapatkan hasil akurasi klasifikasi sebesar 83% dengan precision positif sebesar 87,5% dan precision negatif sebesar 50%. Kesimpulan yang diperoleh dari pengujian akurasi ini adalah bahwa Run-Length Smearing Algorithm dapat digunakan sebagai Optical Character Recognition untuk Pembacaan Angka Meter Air karena tingkat akurasinya yang besar.

5 Kesimpulan



Berdasarkan hasil penelitian yang telah dilakukan dengan menggunakan 90 data sampel data gambar meteran air, bahwa Identifikasi Angka Meter Air Dengan Metode Run-Length Smearing Algorithm berhasil penulis bangun dan Metode RLSA dapat mengenali pola karakter angka pada citra Angka Meter Air, dan berdasarkan tingkat pengukuran dari hasil penelitian yang dilakukan yaitu menunjukkan tingkat keberhasilan identifikasi Angka meteran air yaitu 83%.

6 Daftar Pustaka

- [1] Gunawan dkk, 2014. Penerapan Optical Character Recognition (Ocr) Untuk Pembacaan Meteran Listrik PLN. Informatika Vol. 10 No. 2. Teknik Informatika Fakultas Teknologi Informasi UKDW. Yogyakarta.
- [2] Kenneth E. Kendall, Julie E. Kendall. 2010. Analisis dan Perancangan Sistem. PT Indeks, Jakarta.
- [3] Liliana dkk, 2010. Segmentasi Plat Nomor Kendaraan Dengan Menggunakan Metode Run-Length Smearing Algorithm (RLSA). http://fportfolio.petra.ac.id/user_files/02-030/%5BG-D303-08%5D%20Segmentasi%20Plat%20Nomor%20Kendaraan.pdf. Diakses 23 Oktober 2012.
- [4] Nazruddin Safaat H, 2011, Android (Pemrograman Aplikasi Mobile Smartphone dan Tablet PC Berbasis Android). Informatika, Bandung.
- [5] Priyanto dan Nursari, 2016. Aplikasi Sistem Informasi Pencatatan Meteran Air Menggunakan Qr Code Berbasis Smartphone. Prosiding Seminar Nasional Aplikasi Sains & Teknologi (SNAST). Yogyakarta.
- [6] Putra, Erwin Dwika. 2012. Algoritma Template Matching Pada OCR (Optical Character Recognition). <http://www.academia.edu/30256>

- 516/Algoritma_Template_Matching_Pada_OCR_Optical_Character_Recognition.
Diakses 23 Oktober 2017.
- [7] Pressman, Roger, S. (2001), Software Engineering: A Practitioner's Approach, Fifth Ed. New York, McGraw-Hill Book Company.
- [8] Pakaja, F., Naba, A., dan Purwanto, 2012. Peramalan Penjualan Mobil Menggunakan Jaringan Syaraf Tiruan dan Certainty Factor, Jurnal EECCIS, vol. 6, no. 1.
- [9] Yadav, D., Sanchez, S., & Jorge, M., "Optical Character Recognition for Hindi Language". Journal of Information Processing Systems, Vol.9, No.1, 117-140, 2013.
- [10] Putra, Erwin Dwika, Erzi Hidayat, and Handrie Noprisson. "Model Mobile Positioning System Berbasis Android." Pseudocode 3.2 (2016): 113-121.
- [11] Putra, Erwin Dwika, and Dedy Abdullah. "Deteksi Plat Nomor Kendaraan Roda Dua Menggunakan Metode Edge Detection." INTECOMS: Journal of Information Technology and Computer Science 1.2 (2018): 262-266.

7 Penulis

	Erwin Dwika Putra, M.Kom Fakultas Teknik Program Studi Teknik Informatika Universitas Muhammadiyah Bengkulu
	Marissa Utami, M.Kom Fakultas Teknik Program Studi Sistem Informasi Universitas Muhammadiyah Bengkulu