

IMPLEMENTASI *OPTICAL CHARACTER RECOGNITION* (OCR) UNTUK MENINGKATKAN AKURASI DAN KECEPATAN INPUT DATA DI POSYANDU

Amelia Marshanda Syahputri¹, Budi Harijanto², Cahya Rahmad³

^{1,2,3}Teknologi Informasi, Politeknik Negeri Malang

¹amelia.marshanda20@gmail.com, ²budi.harijanto@polinema.ac.id, ³cahya.rahmad@polinema.ac.id

Abstrak

Posyandu adalah salah satu lembaga kesehatan masyarakat yang membutuhkan pengumpulan dan pemrosesan data yang efisien dan akurat untuk keperluan pemantauan dan evaluasi. Beberapa Posyandu saat ini telah menggunakan sistem informasi untuk mendukung operasionalnya. Namun, proses pengumpulan data di Posyandu saat ini masih dilakukan secara manual dengan mengetikkan informasi ke dalam sistem, yang rentan terhadap kesalahan dan memakan waktu cukup lama. Penelitian ini bertujuan untuk mengembangkan sistem yang mempermudah kader Posyandu dalam proses penginputan data dengan menerapkan teknologi *Optical Character Recognition* (OCR). Teknologi OCR digunakan untuk mengenali karakter dari teks yang ditulis tangan atau dicetak menjadi format digital yang dapat diproses oleh komputer. Pengujian sistem dilakukan dengan menggunakan data dari alat ukur digital untuk mengukur akurasi dan efisiensi waktu pemrosesan. Hasil pengujian menunjukkan bahwa sistem yang dikembangkan mampu mendeteksi karakter pada alat ukur digital dengan rata-rata akurasi sebesar 82% pada jarak 10 cm di dalam ruangan. Selain itu, rata-rata waktu yang dibutuhkan untuk pemrosesan gambar adalah 2,65 detik per gambar. Berdasarkan hasil tersebut, penerapan teknologi OCR dalam sistem informasi Posyandu terbukti dapat meningkatkan efisiensi penginputan data dengan mempercepat proses dan mengurangi potensi kesalahan.

Kata kunci : Akurasi Data, Efisiensi, *Optical Character Recognition*, Posyandu

1. Pendahuluan

Program kesehatan berbasis masyarakat yang disebut Posyandu di Indonesia bertujuan untuk menurunkan angka kematian ibu dan anak serta meningkatkan pertumbuhan dan perkembangan anak (Farmani et al., 2021). Sebagai bagian dari upaya pemerintah untuk mengurangi angka kematian bayi dan angka kematian ibu (AKI/AKB), pengelolaan dan penyelenggaraan Posyandu harus fleksibel, terjangkau, efektif, dan efisien sehingga pemanfaatan seperti ibu hamil dan balita dapat memanfaatkan layanan sebaik mungkin (Dwi Putra et al., 2022).

Proses pengukuran di Posyandu awalnya dilakukan dengan alat ukur konvensional untuk berat badan, tinggi badan, dan lingkar kepala. Seiring dengan perkembangan digitalisasi, alat ukur tersebut kini telah bertransformasi menjadi versi digital yang dapat menampilkan hasil pengukuran secara lebih presisi (Rizani et al., n.d.). Beberapa Posyandu saat ini telah menggunakan sistem informasi untuk mendukung operasionalnya. Meskipun beberapa Posyandu telah mengadopsi sistem informasi untuk mendukung operasionalnya, proses pencatatan data di Posyandu masih mengandalkan input manual, di mana kader harus mengetikkan setiap hasil pengukuran ke dalam sistem informasi. Selain memakan waktu, proses input manual ini juga

membuka peluang terjadinya kesalahan pengetikan yang dapat berdampak serius pada akurasi pemantauan pertumbuhan balita

Tujuan dari penelitian ini adalah sebagai berikut:

- Mengimplementasikan *Optical Character Recognition* (OCR) dengan aplikasi Sistem Informasi Posyandu
- Membantu kader Posyandu dalam melakukan input data dengan lebih mudah
- Mengurangi risiko kesalahan input data di Posyandu

Penelitian ini bertujuan menggunakan metode *Optical Character Recognition* (OCR), khususnya *Tesseract OCR* yang menerapkan algoritma *Long Short-Term Memory* (LSTM). Metode ini diterapkan untuk mengotomatisasi proses pencatatan dengan cara mengenali teks dari hasil pengukuran yang diambil melalui gambar atau dokumen yang dipindai. Dengan teknologi ini, data hasil pengukuran dapat langsung diubah menjadi data digital yang dapat diolah dan disimpan secara otomatis.

Penggunaan OCR dengan metode LSTM diharapkan mampu meningkatkan efisiensi proses pencatatan, mengurangi risiko kesalahan input, serta memastikan akurasi dan keamanan data dalam sistem informasi Posyandu. Dengan demikian, penelitian ini bertujuan untuk mendukung pengelolaan Posyandu

yang lebih efektif dan efisien, serta berkontribusi pada pemantauan kesehatan dan perkembangan balita yang lebih baik.

2. Tinjauan Pustaka

2.1 Studi Literatur

Beberapa penelitian terdahulu yang memiliki keterkaitan dengan penelitian ini, untuk menentukan posisi penelitian. Adapun penelitian terdahulu yang peneliti maksud adalah:

Penelitian oleh Setiawan (2017) dalam Jurnal Sistem dan Teknologi Informasi menunjukkan bahwa implementasi OCR pada aplikasi penerjemah bahasa Indonesia ke bahasa Inggris berbasis Android dapat menangkap teks dengan baik, meskipun memiliki kesulitan dengan huruf kapital, tanda baca, dan font tertentu seperti *Segoe Script* dan *Lucida Handwriting* (Setiawan A., 2017).

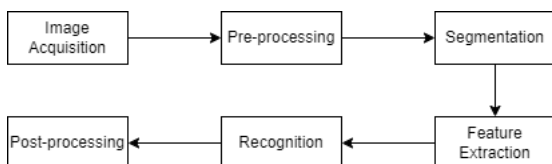
Ibnutama (2019) meningkatkan akurasi deteksi plat nomor kendaraan dengan memodifikasi metode *template matching* pada OCR, termasuk perubahan nilai ambang dan penambahan tahapan dilasi, yang meningkatkan deteksi sebesar 4,44% pada 35 sampel (Ibnutama et al., 2019).

Sanjaya dan Sugiarto (2022) menggunakan *Deep Learning Convolutional Neural Network* (CNN) untuk sistem pemantauan meteran air berbasis OCR, yang menunjukkan akurasi tinggi pada angka meteran empat digit awal, tetapi mengalami kegagalan tinggi pada dua digit akhir (Sanjaya & Sugiarto, 2022).

2.2 Optical Character Recognition (OCR)

Optical Character Recognition adalah teknologi yang mengubah teks tertulis tangan atau cetak menjadi format digital yang dapat diproses oleh komputer, memungkinkan konversi dokumen fisik menjadi format digital yang bisa diedit dan dicari. Teknologi ini sangat berguna untuk pengolahan dokumen, seperti pengarsipan, pengolahan data, dan pengenalan plat nomor kendaraan (AISSMS Institute of Information Technology & Institute of Electrical and Electronics Engineers, 2020).

Proses pengenalan karakter yang dilakukan OCR adalah seperti Gambar 1 berikut:



Gambar 1. Alur Diagram OCR

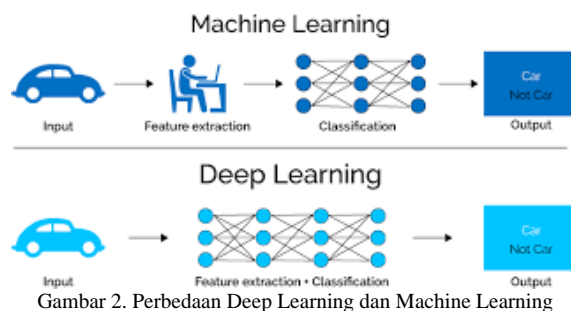
2.3 Tesseract OCR

Tesseract OCR adalah mesin OCR sumber terbuka yang digunakan untuk mengenali teks pada gambar atau dokumen yang dipindai, mengubahnya

menjadi format yang dapat dibaca oleh mesin. Mesin ini mendukung berbagai bahasa, termasuk bahasa Indonesia, dan dapat diakses secara gratis dari situs web resminya (Ramdhani et al., 2021). *Tesseract* OCR bekerja dengan mengenali pola pada gambar atau dokumen untuk mengidentifikasi karakter atau huruf, kemudian mengubahnya menjadi teks digital menggunakan algoritma pengenalan karakter yang kompleks. Algoritma ini terus ditingkatkan untuk meningkatkan akurasi dan mesin dapat disesuaikan untuk mengenali teks dalam berbagai bahasa menggunakan model bahasa yang telah deprogram sebelumnya.

2.4 Deep Learning

Deep learning adalah teknik kecerdasan buatan yang meniru cara otak manusia dalam memproses informasi. Secara umum, algoritma *deep learning* mirip dengan algoritma *machine learning*, namun pada *deep learning*, proses ekstraksi fitur dan klasifikasi dilakukan secara bersamaan hingga menghasilkan output (Ramba, 2020). Perbandingan antara algoritma *deep learning* dan *machine learning* dapat dilihat pada Gambar 2.



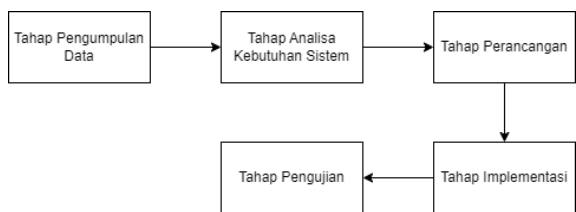
Gambar 2. Perbedaan Deep Learning dan Machine Learning

2.5 Long Short Term Memory (LSTM)

LSTM (*Long Short-Term Memory*) adalah pengembangan dari arsitektur RNN (*Recurrent Neural Network*) yang diperkenalkan oleh Hochreiter & Schmidhuber. Berbeda dengan RNN, LSTM mampu mengatasi kelemahan RNN dalam mengingat informasi jangka panjang karena menggunakan *memory cells* dan *gate units*. LSTM dapat menyimpan atau membuang data melalui *forget gate*, memperbarui data melalui *input gate*, dan mengeluarkan data melalui *output gate*. Proses ini memungkinkan LSTM untuk mengingat informasi penting dalam jangka panjang dan melacak berbagai atribut teks yang sedang diproses (Arief et al., 2023).

3. Metodologi Penelitian

Metodologi penelitian yang diterapkan dalam penelitian ini terdiri dari lima tahap utama, seperti yang diilustrasikan pada Gambar 3. Tahapan tersebut meliputi pengumpulan data, Analisa kebutuhan sistem, perancangan sistem, implementasi, dan pengujian.



Gambar 3 Alur Penelitian

Tahap pengumpulan data dilaksanakan melalui observasi lapangan untuk memperoleh data primer. Observasi ini mencakup dokumentasi alur kerja kader Posyandu, inventarisasi peralatan yang digunakan, serta identifikasi permasalahan operasional yang dihadapi dalam proses pencatatan data kesehatan balita.

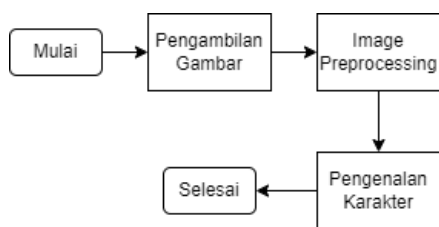
Berdasarkan data dan informasi yang terkumpul, penelitian berlanjut ke tahap analisis kebutuhan sistem. Pada tahap ini, dilakukan analisis mendalam terhadap kebutuhan fungsional dan non-fungsional sistem, dengan mempertimbangkan aspek teknis maupun operasional. Langkah ini diambil agar dapat merancang solusi yang tepat untuk mengatasi kendala-kendala yang telah teridentifikasi.

Tahap perancangan meliputi desain arsitektur sistem secara detail, termasuk perancangan basis data, antarmuka pengguna, dan alur proses sistem. Perancangan dilakukan dengan mempertimbangkan kemudahan penggunaan sistem oleh kader Posyandu.

Implementasi sistem merupakan tahap realisasi dari rancangan yang telah dibuat, dengan mengintegrasikan teknologi OCR ke dalam sistem informasi Posyandu.

Tahap final dari metodologi penelitian ini adalah pengujian sistem, yang dilakukan untuk mengevaluasi kinerja dan efektivitas sistem. Pengujian meliputi evaluasi akurasi OCR dalam membaca data dan pengukuran efisiensi waktu pemrosesan. Hasil pengujian dianalisis untuk menilai tingkat keberhasilan sistem dalam mengatasi permasalahan yang telah diidentifikasi sebelumnya.

Alur pengolahan data yang digunakan dalam penelitian ini terdiri dari 3 tahap yang berurutan ditunjukkan pada Gambar 4.



Gambar 4. Alur Pengolahan Data

Pengolahan data dalam sistem dimulai dengan pengambilan gambar hasil pengukuran melalui kamera pengguna, kemudian disimpan dan dilakukan proses *image preprocessing* untuk memudahkan pengenalan tiap karakter. Selanjutnya dilakukan pengenalan karakter dan mengubahnya menjadi teks menggunakan mesin *Tesseract OCR*, yang kemudian disimpan dalam *database* sebagai data kesehatan

peserta Posyandu. Data ini dapat diakses oleh kader Posyandu untuk memantau pertumbuhan dan kesehatan peserta, sehingga sistem ini mempermudah pemantauan dan pengelolaan data kesehatan di Posyandu.

Parameter yang akan digunakan dalam pengujian meliputi kondisi pencahayaan dan jarak saat pengambilan gambar untuk mengukur tingkat akurasi pengenalan karakter. Rencana pengujian dapat dilihat pada Tabel 1 dan Tabel 2.

Tabel 1. Rancangan Pengujian pada Intensitas Cahaya

| Test Case | Deskripsi | Hasil Harapan |
|---------------|--|--|
| Luar Ruangan | Pengambilan gambar dengan cahaya luar ruangan | OCR dapat mengenali dan mengekstrak teks dengan cepat dan akurat pada kondisi cahaya terang |
| Dalam Ruangan | Pengambilan gambar dengan cahaya dalam ruangan | OCR dapat mengenali dan mengekstrak teks dengan baik meskipun dalam kondisi cahaya yang kurang optimal |

Tabel 2. Rancangan Pengujian Terhadap Jarak

| Test Case | Deskripsi | Hasil Harapan |
|------------------------|--|--|
| Jarak Dekat (5 cm) | Pengambilan gambar dari jarak dekat | OCR dapat mengenali dan mengekstrak teks dengan baik pada jarak dekat |
| Jarak menengah (10 cm) | Pengambilan gambar dari jarak menengah | OCR memberikan hasil yang memadai pada jarak pengambilan gambar yang umum digunakan |
| Jarak Jauh (15 cm) | Pengambilan gambar dari jarak jauh | OCR dapat memberikan hasil yang akurat pada jarak pengambilan gambar yang cukup jauh |

Selain itu, untuk mengukur efisiensi atau kecepatan, waktu yang dibutuhkan dalam pengolahan data hingga data siap disimpan di database akan dijadikan sebagai parameter. Untuk perancangan pengujian dapat dilihat pada Tabel 3.

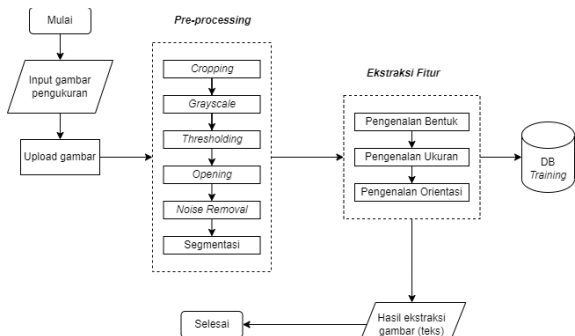
Tabel 3. Rancangan Pengujian Terhadap Kecepatan Performa OCR

| Gambar | Waktu Mulai | Waktu Selesai | Total Waktu |
|--|----------------------------------|----------------------------------|-----------------------------|
| Gambar hasil pengukuran alat ukur digital dengan format jpg/png Dst. | Waktu dimulainya pemrosesan data | Waktu selesainya pemrosesan data | Waktu selesai – waktu mulai |

4. Hasil dan Pembahasan

Proses pengujian dilakukan sesuai dengan alur perancangan pengujian pada Gambar 5. Pengujian dilakukan pada gambar hasil pengukuran yang

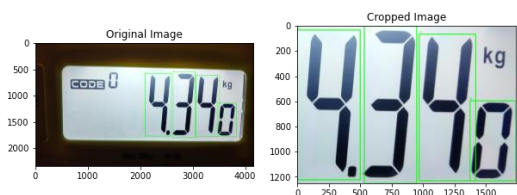
selanjutnya akan menjadi data yang dapat diolah atau disimpan oleh komputer.



Gambar 5. Alur Pengujian

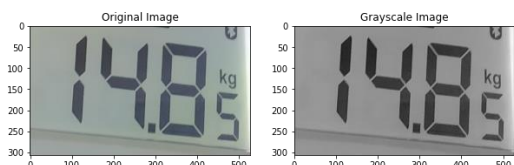
4.1 Image Preprocessing

Pengolahan citra pada proses *image preprocessing* meliputi *cropping*, *grayscale*, *thresholding*, *opening*, *noise removal*, dan *segmentasi*. Proses *cropping* dilakukan untuk membantu menghilangkan area yang tidak relevan dan memungkinkan fokus pada objek atau area penting dalam gambar. Metode yang digunakan dalam proses *cropping* ini adalah pemilihan koordinat atau batasan (x, y) dari area objek yang menjadi acuan area yang akan dipotong seperti pada Gambar 6.



Gambar 6. Hasil Cropping

Proses *grayscale* dilakukan untuk mengurangi kompleksitas data dan mempercepat proses pengolahan gambar. Dalam gambar *grayscale*, setiap piksel direpresentasikan oleh satu saluran warna saja, yaitu tingkat kecerahan (intensitas) dari hitam ke putih. Implementasi dapat dilihat seperti pada Gambar 7.



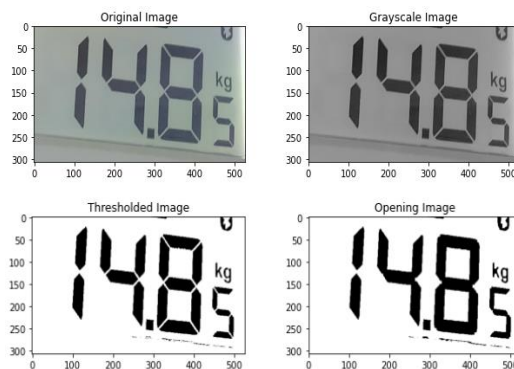
Gambar 7. Hasil Grayscale

Thresholding adalah teknik yang digunakan untuk mengonversi gambar grayscale menjadi gambar biner. Seperti pada Gambar 8, setiap piksel hanya memiliki nilai hitam (0) atau putih (1).



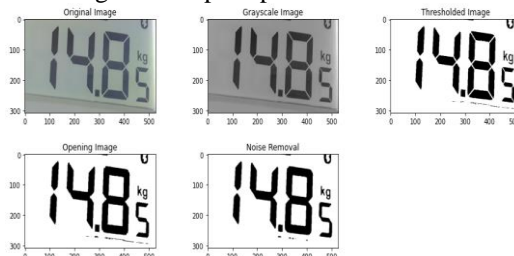
Gambar 8. Hasil Thresholding

Selanjutnya proses *opening* yang berguna untuk menghilangkan objek kecil atau noise dari gambar, terutama yang tidak sesuai dengan struktur atau ukuran tertentu, tanpa mengubah bentuk atau ukuran objek yang lebih besar dalam gambar. Hasil implementasi *opening* seperti pada Gambar 9.



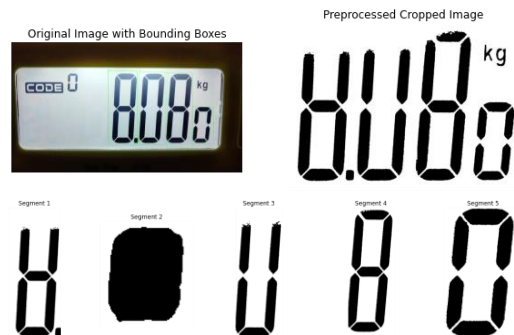
Gambar 9. Hasil Opening

Noise pada gambar dapat berasal dari berbagai sumber, seperti kondisi pencahayaan yang buruk, sensor kamera yang tidak sempurna, atau proses transmisi gambar. Langkah penghapusan *noise* bertujuan untuk mengurangi atau menghilangkan gangguan yang tidak diinginkan dari gambar, sehingga memperbaiki kualitas gambar seperti pada Gambar 10 berikut.



Gambar 10. Hasil Noise Removal

Dalam konteks pengenalan teks menggunakan OCR, segmentasi digunakan untuk memisahkan teks dari latar belakang dan membedakan karakter atau kata satu sama lain. Proses segmentasi gambar akan dibagi berdasarkan jumlah karakter yang terbaca oleh OCR. Hasil segmentasi divisualisasikan pada Gambar 11.



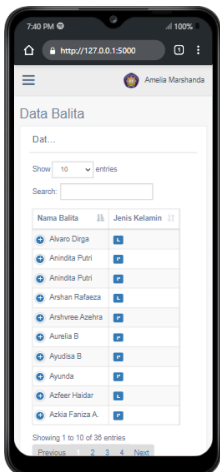
Gambar 11. Hasil Segmentasi

4.2 Feature Extraction

Ekstraksi fitur pada proses ini adalah mengambil fitur-fitur penting dari gambar teks untuk mengenali karakter di dalamnya. Proses ini membantu sistem OCR membedakan karakter yang mirip, meningkatkan akurasi, dan menangani variasi dalam gambar. Dalam OCR seperti *Tesseract*, *feature extraction* berperan penting dalam mengenali teks dengan lebih akurat, bahkan dalam kondisi gambar yang kurang ideal.

4.3 Implementasi Sistem

Sistem informasi Posyandu yang dibangun untuk penelitian ini menggunakan framework Flask yang memungkinkan pengembangan aplikasi web dengan cepat dan efisien. Pada saat awal mengakses sistem, data balita dari Posyandu akan ditampilkan untuk mengisi data pengukuran, pengguna dapat langsung menekan nama balita yang akan diinput data pengukurannya. Hasil implementasi seperti pada Gambar 12.



Gambar 12. Tampilan Halaman Data Balita

Gambar 13 merupakan halaman yang digunakan untuk mengambil hasil pengukuran dari alat ukur digital. Pengguna akan mengambil gambar hasil pengukuran, kemudian setelah hasilnya muncul, data tersebut dapat langsung disimpan di dalam *database*.



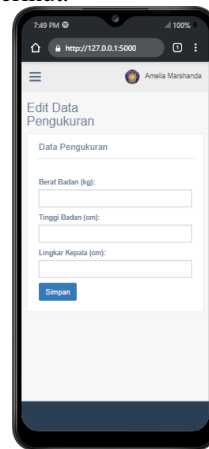
Gambar 13. Tampilan Halaman Ambil Gambar

Setelah mendapatkan data hasil pengukuran, data tersebut akan ditampilkan pada halaman data pengukuran sebagai bukti telah tersimpan di *database*. Tampilan hasil dari data pengukuran terlihat seperti pada Gambar 14.



Gambar 14. Tampilan Halaman Data Pengukuran

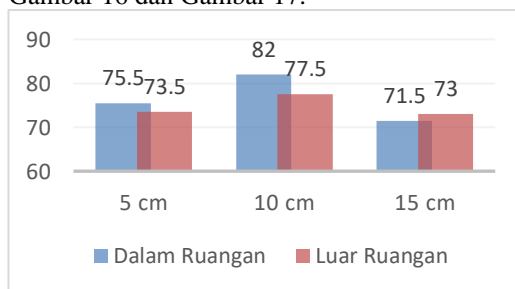
Pada sistem ini juga dilengkapi dengan halaman input data manual. Halaman ini dapat digunakan apabila ingin menginputkan data hasil pengukuran secara manual melalui form yang disediakan seperti pada Gambar 15 berikut.



Gambar 15. Tampilan Halaman Input Data Manual

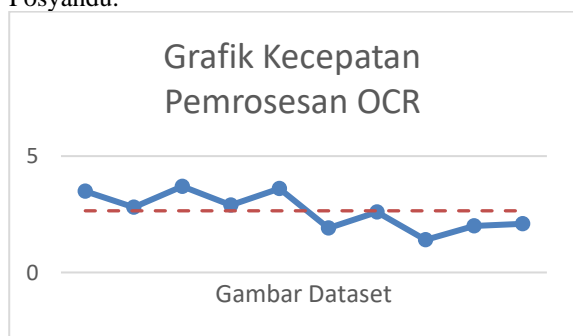
4.4 Hasil Pengujian

Jumlah data yang digunakan dalam penelitian ini adalah 70 sampel yang terdiri dari 30 sampel di dalam ruangan, 30 sampel di luar ruangan, dan 10 sampel untuk pengujian kecepatan performa OCR. Hasil pengujian dapat divisualisasikan seperti pada Gambar 16 dan Gambar 17.



Gambar 16. Diagram Perbandingan Hasil Pengujian

Pada diagram diatas, terlihat bahwa tingkat akurasi tertinggi didapat pada saat jarak pengambilan gambar 10 cm di dalam ruangan sebesar 82%. Sedangkan pada pengujian terhadap efisiensi, terlihat pada grafik bahwa waktu yang dibutuhkan untuk mengelola data relative stabil dengan rata-rata 2,65 detik per gambar. Dari hasil pengujian ini dinilai bahwa sistem yang dikembangkan mampu meningkatkan akurasi dan efisiensi kerja kader Posyandu.



Gambar 17 Grafik Hasil Pengujian Kecepatan OCR

5. Kesimpulan dan Saran

Implementasi *Optical Character Recognition* (OCR) secara signifikan membantu kader Posyandu dalam pengukuran dan pencatatan data bayi, meningkatkan akurasi dan kecepatan input data. Pengujian menunjukkan bahwa kinerja *Tesseract* OCR dipengaruhi oleh jarak pengambilan gambar dan kondisi pencahayaan. Akurasi tertinggi (79,7%) dicapai pada jarak 10 cm, sedangkan terendah (72,2%) pada jarak 15 cm. pencahayaan dalam ruangan memberikan akurasi terbaik (76,3%) dibandingkan luar ruangan, yang rentan terhadap ketidakseimbangan cahaya. Meskipun efisien dengan rata-rata waktu 2,65 detik per gambar, OCR masih memiliki kelemahan dalam mengenali karakter tanda baca. Namun, teknologi ini tetap efektif dalam mempercepat penginputan data di Posyandu.

Pada penelitian ini, hasil deteksi karakter menggunakan *Tesseract* OCR pada alat ukur digital menunjukkan akurasi yang cukup baik, namun masih terdapat kelemahan dalam mengenali karakter tanda baca. Untuk mengatasi masalah ini, beberapa solusi yang diusulkan adalah melakukan pelatihan tambahan pada *Tesseract* OCR dengan dataset yang lebih beragam untuk meningkatkan kemampuan pengenalan karakter, terutama dalam kondisi yang kurang ideal, serta melatih *Tesseract* OCR menggunakan dataset khusus yang mencakup berbagai tanda baca untuk meningkatkan akurasi deteksi.

Daftar Pustaka:

AISSMS Institute of Information Technology, & Institute of Electrical and Electronics

Engineers. (2020). *2020 International Conference on Emerging Smart Computing and Informatics (ESCI): AISSMS Institute of Information Technology, Pune, India. Mar 12-14, 2020.*

Arief, B., Kholifatullah, H., & Prihanto, A. (2023). Penerapan Metode Long Short Term Memory Untuk Klasifikasi Pada Hate Speech. *Journal of Informatics and Computer Science*, 04.

Dwi Putra, A., Pratiwi, T., & Asharudin, F. (2022). Sistem Informasi Posyandu Dusun Pelemgede Desa Sodo Kecamatan Paliyan Kabupaten Gunungkidul. In *Information Sistem Journal (INFOS)* / (Vol. 5, Issue 1).

Farmani, P. I., Adiputra, I. N. M., & Laksmi, P. A. (2021). Perancangan Sistem Informasi Posyandu Sebagai Upaya Digitalisasi Data Posyandu di UPTD Puskesmas II Dinas Kesehatan Kecamatan Denpasar Timur. *Indonesian of Health Information Management Journal (INOHIM)*, 9(2), 115–126. <https://doi.org/10.47007/inohim.v9i2.311>

Ibnutama, K., Panjaitan, Z., & Ginting, E. F. (2019). J-SISKO TECH Jurnal Teknologi Sistem Informasi dan Sistem Komputer TGD Modifikasi Metode Template Matching pada OCR Untuk Meningkatkan Akurasi Deteksi Plat Nomor Kendaraan. v, 21(2), 21–29.

Ramba, L. S. (2020). Design Of A Voice Controlled Home Automation System Using Deep Learning Convolutional Neural Network (DL-CNN). *Telekontran: Jurnal Ilmiah Telekomunikasi, Kendali Dan Elektronika Terapan*, 8(1), 57–73. <https://doi.org/10.34010/telekontran.v8i1.3078>

Ramdhani, T. W., Budi, I., & Purwandari, B. (n.d.). Optical Character Recognition Engines Performance Comparison in Information Extraction. In *IJACSA International Journal of Advanced Computer Science and Applications* (Vol. 12, Issue 8). www.ijacsa.thesai.org

Rizani, F., Informatika, J., Komputer, I., Almuslim, U., Almuslim, J., & Cut Bireuen -Aceh, P. (n.d.). Sistem Informasi Posyandu Kabupaten Bireuen Berbasis Authorization Managemant System. *Jurnal TIKA Fakultas Ilmu Komputer Universitas Almuslim*, 7(1).

Sanjaya, I., & Sugiarto, I. (2022). Sistem Pemantauan Meteran Air Berbasis Optical Character Recognition. In *Jurnal Teknik Elektro* (Vol. 15, Issue 2). <https://embedded-lab.com/blog/high-voltage-seven-segment->

Setiawan A, S. H. B. A. (2017). Implementasi Optical Character Recognition (OCR) pada Mesin Penerjemah Bahasa Indonesia ke Bahasa Inggris. *Jurnal Sistem Dan Teknologi Informasi (JUSTIN)*, 5.