BAB II

TINJAUAN LITERATUR

2.1 Penelitian Terkait

Penelitian mengenai *Optical Character Recognition* (OCR) telah banyak dilakukan untuk mengekstraksi informasi teks dari berbagai jenis dokumen digital. Berbagai penelitian menunjukkan bahwa teknologi OCR mampu mengubah teks pada citra menjadi data digital sehingga dapat dimanfaatkan dalam proses digitalisasi dokumen, pengelolaan arsip, serta ekstraksi informasi secara otomatis. Namun demikian, hasil pengenalan teks masih dipengaruhi oleh kualitas citra, tata letak dokumen, jenis huruf, serta kondisi lingkungan saat pengambilan gambar. Oleh karena itu, evaluasi terhadap kinerja sistem OCR masih terus dilakukan pada berbagai jenis dokumen untuk memperoleh metode yang paling sesuai (Darpito et al., 2025).

Beberapa penelitian telah memanfaatkan EasyOCR sebagai sistem pengenalan teks karena bersifat *open source*, mudah diimplementasikan, dan mampu mengenali berbagai bahasa. Penelitian oleh (Sukarna et al., 2025) menunjukkan bahwa EasyOCR mampu menghasilkan proses ekstraksi teks yang baik pada dokumen digital, namun performanya dipengaruhi oleh kualitas citra dan karakteristik dokumen. Penelitian lain oleh (Hanif et al., 2026) juga menunjukkan bahwa EasyOCR memiliki tingkat kesalahan yang berbeda ketika dibandingkan dengan sistem OCR lainnya sehingga diperlukan evaluasi menggunakan parameter yang sama.

Selain EasyOCR, PaddleOCR juga banyak digunakan karena menggabungkan proses deteksi dan pengenalan teks berbasis *deep learning* dengan waktu pemrosesan yang relatif cepat (Li et al., 2022). Penelitian oleh (Nazeem et al., 2024) menunjukkan bahwa PaddleOCR memiliki kemampuan yang baik dalam mengenali teks pada berbagai kondisi citra. Meskipun demikian, beberapa penelitian menyimpulkan bahwa belum terdapat metode OCR yang selalu memberikan hasil terbaik pada seluruh jenis dokumen. Oleh karena itu, penelitian ini melakukan analisis komparatif antara EasyOCR dan PaddleOCR menggunakan objek berupa dokumen ijazah dengan parameter evaluasi Character Error Rate (CER), Word Error Rate (WER), dan waktu pemrosesan.

Berdasarkan beberapa penelitian terdahulu, dapat diketahui bahwa teknologi OCR telah banyak diterapkan pada berbagai jenis dokumen dengan menggunakan sistem seperti EasyOCR dan PaddleOCR. Hasil penelitian menunjukkan bahwa setiap sistem memiliki karakteristik dan tingkat kinerja yang berbeda bergantung pada jenis dokumen, kualitas citra, serta metode evaluasi yang digunakan. Namun, penelitian yang secara khusus membandingkan kinerja EasyOCR dan PaddleOCR pada dokumen ijazah masih terbatas. Selain itu, belum banyak penelitian yang melakukan evaluasi menggunakan dataset dokumen ijazah yang sama dengan parameter Character Error Rate (CER), Word Error Rate (WER), dan waktu pemrosesan. Oleh karena itu, penelitian ini dilakukan untuk mengisi kesenjangan tersebut melalui analisis komparatif EasyOCR dan PaddleOCR dalam pengenalan teks pada dokumen ijazah sehingga diperoleh informasi mengenai sistem OCR yang lebih sesuai berdasarkan hasil evaluasi yang dilakukan.

Tabel 2. 1 Tinjauan Pustaka

Nama Peneliti	Judul Penelitian	Hasil Penelitian	Perbedaan dengan Penelitian yang Akan Dilakukan
Darpito et al. (2025)	Perbandingan Unjuk Kerja Library Optical Character Recognition (OCR) dalam Pengenalan Teks pada Dokumen Digital	Penelitian membandingkan beberapa library OCR dalam pengenalan teks pada dokumen digital menggunakan parameter evaluasi seperti Character Error Rate (CER), Word Error Rate (WER), dan waktu pemrosesan. Hasil penelitian menunjukkan bahwa setiap library OCR memiliki tingkat akurasi dan efisiensi yang berbeda dalam mengenali teks pada dokumen digital.	Penelitian ini membandingkan EasyOCR dan PaddleOCR secara khusus pada dokumen ijazah menggunakan dataset yang sama, ground truth, serta parameter CER, WER, dan waktu pemrosesan.
Sukarna et al. (2025)	Studi Komparatif Model OCR Berbasis AI untuk Dokumen Cetak dan Tulisan Tangan	Penelitian membandingkan beberapa model OCR berbasis AI pada dokumen cetak dan tulisan tangan. Hasil penelitian menunjukkan bahwa karakteristik dokumen sangat memengaruhi performa masing-masing model OCR.	Penelitian ini hanya membandingkan EasyOCR dan PaddleOCR dengan objek penelitian berupa dokumen ijazah, bukan dokumen cetak dan tulisan tangan.

Nama Peneliti	Judul Penelitian	Hasil Penelitian	Perbedaan dengan Penelitian yang Akan Dilakukan
Hanif et al. (2026)	Perbandingan Kinerja EasyOCR dan Tesseract dalam Ekstraksi Teks pada Citra Digital	Penelitian membandingkan kinerja EasyOCR dan Tesseract dalam melakukan ekstraksi teks pada citra digital menggunakan parameter CER, WER, dan waktu pemrosesan. Hasil penelitian menunjukkan adanya perbedaan tingkat kesalahan dan waktu pemrosesan antara kedua sistem OCR.	Penelitian ini membandingkan EasyOCR dengan PaddleOCR, bukan EasyOCR dengan Tesseract, serta menggunakan objek penelitian berupa dokumen ijazah.
Li et al. (2022)	*PP-OCRv3: More Attempts for the Improvement of Ultra Lightweight OCR Systemaugmentation*	Penelitian mengembangkan PP-OCRv3 dengan peningkatan pada model deteksi dan pengenalan teks sehingga mampu meningkatkan akurasi tanpa mengurangi efisiensi inferensi.	Penelitian ini tidak mengembangkan model PaddleOCR, tetapi menggunakan PaddleOCR sebagai salah satu sistem OCR yang dibandingkan dengan EasyOCR pada dokumen ijazah.
Nazeem et al. (2024)	Comparative Analysis of OCR Libraries for Text Recognition	Penelitian membandingkan beberapa library OCR dan menunjukkan bahwa setiap library memiliki karakteristik serta performa yang berbeda pada berbagai kondisi citra dan jenis	Penelitian ini berfokus pada perbandingan EasyOCR dan PaddleOCR menggunakan dokumen ijazah dengan evaluasi berdasarkan CER, WER, dan waktu pemrosesan.

Nama Peneliti	Judul Penelitian	Hasil Penelitian	Perbedaan dengan Penelitian yang Akan Dilakukan
		dokumen.	

2.2 Pengolahan Citra Digital

Pengolahan citra digital (*Digital Image Processing*) merupakan suatu bidang ilmu yang mempelajari teknik untuk memperoleh, memanipulasi, menganalisis, dan meningkatkan kualitas citra digital menggunakan bantuan komputer. Tujuan utama pengolahan citra digital adalah menghasilkan citra yang memiliki kualitas lebih baik sehingga informasi yang terkandung di dalamnya dapat ditampilkan, dianalisis, atau diinterpretasikan dengan lebih mudah. Berbagai teknik dalam pengolahan citra digital meliputi peningkatan kualitas citra (*image enhancement*), pengurangan derau (*noise reduction*), perbaikan kontras, segmentasi, serta transformasi citra. Seiring perkembangan teknologi, pengolahan citra digital telah banyak dimanfaatkan dalam berbagai bidang, seperti kesehatan, industri, keamanan, pertanian, penginderaan jauh, dan sistem otomatis yang memanfaatkan citra sebagai sumber informasi (Ashillah et al., 2025).

2.3 OCR

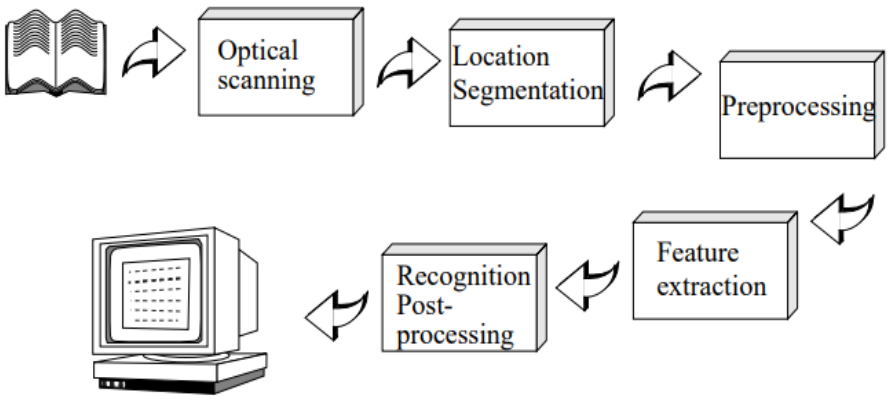
Optical Character Recognition (OCR) merupakan teknologi yang digunakan untuk mengenali karakter atau teks yang terdapat pada citra dan mengubahnya menjadi data digital yang dapat diproses oleh komputer. Citra yang digunakan dapat berupa dokumen hasil pemindaian, foto dokumen, maupun berbagai bentuk

citra digital lainnya. Dengan adanya teknologi OCR, informasi teks yang sebelumnya hanya tersedia dalam bentuk gambar dapat diekstraksi sehingga lebih mudah untuk disimpan, dicari, dan diproses lebih lanjut.

Secara umum, proses OCR terdiri atas dua tahapan utama, yaitu deteksi teks dan pengenalan teks. Tahap deteksi teks bertujuan untuk menemukan lokasi teks yang terdapat pada citra, sedangkan tahap pengenalan teks bertujuan untuk mengidentifikasi karakter atau kata yang berada pada area teks tersebut. Hasil dari proses tersebut kemudian diubah menjadi teks digital. Dalam perkembangannya, sistem OCR modern telah banyak memanfaatkan metode *deep learning* untuk meningkatkan kemampuan dalam mendeteksi dan mengenali teks pada berbagai kondisi citra dan bentuk dokumen (Wang et al., 2021).

Penerapan OCR dipengaruhi oleh karakteristik citra dan dokumen yang diproses. Faktor seperti kualitas citra, resolusi, pencahayaan, kemiringan, jenis huruf, tata letak teks, serta keberadaan noise dapat memengaruhi hasil pengenalan teks. Oleh karena itu, sistem OCR perlu diuji menggunakan dataset yang sesuai dengan objek penelitian untuk mengetahui kemampuan pengenalan teks secara objektif. Dalam penelitian ini, teknologi OCR digunakan untuk mengenali teks pada dokumen ijazah menggunakan dua sistem, yaitu EasyOCR dan PaddleOCR, yang selanjutnya dibandingkan berdasarkan hasil pengenalan teks, tingkat kesalahan karakter dan kata, serta waktu pemrosesan.

Secara umum, tahapan proses OCR dalam penelitian ini dapat digambarkan sebagai berikut:



Gambar 2. 1 Components of an OCR-system (Bezerra & de Oliveira, 2013).

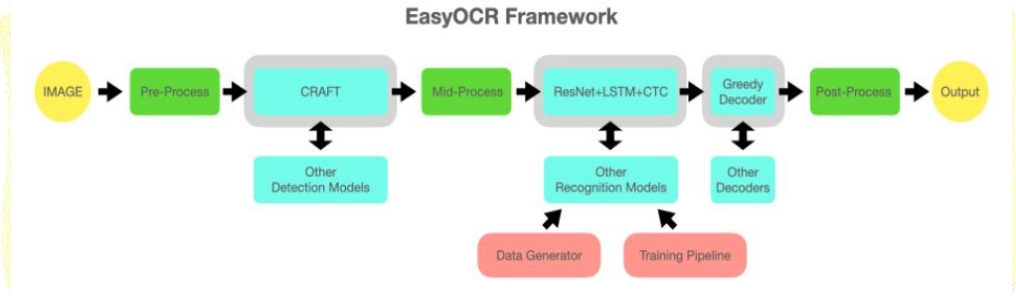
Berdasarkan uraian tersebut, OCR dapat digunakan sebagai teknologi untuk mengubah informasi teks pada dokumen ijazah dalam bentuk citra menjadi teks digital. Namun, karena setiap sistem OCR memiliki karakteristik dan kemampuan yang berbeda, diperlukan perbandingan menggunakan kondisi pengujian dan dataset yang sama untuk mengetahui sistem yang lebih sesuai dalam mengenali teks pada dokumen ijazah.

2.4 EasyOCR

EasyOCR merupakan sistem *Optical Character Recognition* (OCR) berbasis *deep learning* yang digunakan untuk mendeteksi dan mengenali teks pada citra secara otomatis. EasyOCR merupakan perangkat lunak *open-source* yang mendukung berbagai bahasa dan jenis tulisan. Sistem ini dapat menerima citra sebagai masukan dan menghasilkan teks yang berhasil dikenali beserta lokasi teks pada citra dan nilai tingkat keyakinan (*confidence score*) dari hasil pengenalan. Selain itu, EasyOCR mampu mengenali teks pada berbagai jenis dokumen seperti kartu identitas, sertifikat, faktur, maupun dokumen akademik (Hanif et al., 2026; Sukarna et al., 2025)

2.4.1 Arsitektur EasyOCR

Secara umum, EasyOCR memiliki dua komponen utama, yaitu text detection dan text recognition. Pada tahap *text detection*, EasyOCR menggunakan metode CRAFT (Character Region Awareness for Text Detection) untuk menemukan lokasi teks pada citra. Selanjutnya, area teks yang telah terdeteksi diproses pada tahap *text recognition* menggunakan arsitektur CRNN (Convolutional Recurrent Neural Network). Arsitektur CRNN terdiri atas bagian ekstraksi fitur citra, pemodelan urutan teks menggunakan jaringan recurrent, serta proses *decoding* untuk menghasilkan karakter atau teks.



Gambar 2. 2 Arsitektur EasyOCR (Hanif et al., 2026)

2.4.2 Alur Kerja EasyOCR

Proses kerja EasyOCR dimulai dengan memasukkan citra dokumen ke dalam sistem. Citra tersebut kemudian diproses oleh model deteksi untuk menemukan area yang mengandung teks. Setelah area teks berhasil ditemukan, bagian tersebut diproses oleh model pengenalan untuk membaca karakter dan menyusunnya menjadi teks. Hasil proses EasyOCR dapat berupa teks yang dikenali, koordinat lokasi teks pada citra, serta nilai *confidence score*.

Secara sederhana, alur kerja EasyOCR dapat digambarkan sebagai berikut:

1. Citra Dokumen

Merupakan gambar masukan yang akan diproses oleh EasyOCR, misalnya citra atau foto dokumen ijazah yang berisi berbagai informasi teks.

2. Deteksi Teks (CRAFT)

   Pada tahap ini, metode CRAFT (*Character Region Awareness for Text Detection*) digunakan untuk menemukan lokasi teks pada citra. Sistem mendeteksi area yang mengandung karakter atau teks, seperti nama, nomor ijazah, program studi, dan informasi lainnya.

3. Area Teks

   Hasil deteksi CRAFT berupa wilayah atau kotak pembatas (*bounding box*) yang menunjukkan lokasi teks pada dokumen. Area teks tersebut kemudian dipisahkan atau dipotong untuk diproses pada tahap berikutnya.

4. Pengenalan Teks (CRNN)

   Area teks yang telah terdeteksi diproses menggunakan CRNN (*Convolutional Recurrent Neural Network*) untuk mengenali karakter dan menyusunnya berdasarkan urutan sehingga menghasilkan teks.

5. Teks Digital

   Hasil akhir dari proses tersebut berupa teks digital yang telah dikenali oleh EasyOCR. Teks tersebut dapat ditampilkan, disimpan, atau digunakan untuk proses evaluasi terhadap *ground truth*.

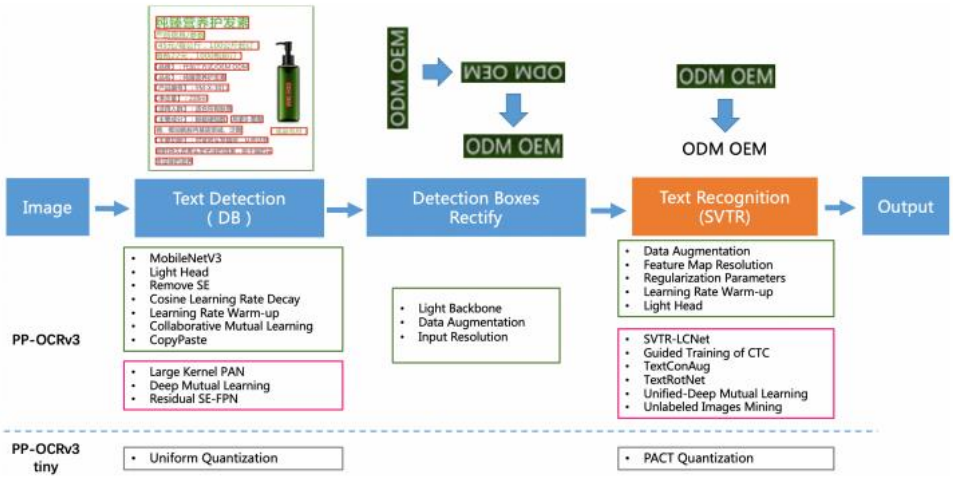
2.5 PaddleOCR

PaddleOCR merupakan *toolkit Optical Character Recognition* (OCR) berbasis *deep learning* yang digunakan untuk mendeteksi dan mengenali teks pada citra secara otomatis. PaddleOCR dikembangkan sebagai sistem OCR yang dapat

digunakan untuk berbagai kebutuhan pengenalan teks pada dokumen dan citra digital. Sistem ini mengintegrasikan proses deteksi teks dan pengenalan teks dalam satu rangkaian pemrosesan sehingga citra yang berisi teks dapat diubah menjadi data teks digital. PaddleOCR menggunakan rangkaian model yang dikenal sebagai PP-OCR (*Practical and Precise OCR*). PP-OCR dikembangkan untuk menghasilkan sistem OCR yang memiliki kemampuan pengenalan teks dengan mempertimbangkan efisiensi pemrosesan. Pengembangan PP-OCR dilakukan melalui beberapa versi, seperti PP-OCRv2 dan PP-OCRv3, yang mengalami peningkatan pada komponen deteksi dan pengenalan teks (Du et al., 2021; Li et al., 2022).

2.5.1 Arsitektur PaddleOCR

Secara umum, arsitektur PaddleOCR terdiri atas beberapa komponen utama, yaitu text detection, direction classification, dan text recognition. *Text detection* digunakan untuk menemukan lokasi teks pada citra. Setelah area teks ditemukan, *direction classification* digunakan untuk menentukan orientasi teks sehingga proses pengenalan dapat dilakukan dengan lebih tepat. Selanjutnya, *text recognition* digunakan untuk mengenali karakter pada area teks dan mengubahnya menjadi teks digital (Du et al., 2021).



Gambar 2. 3 Arsitektur PaddleOCR (Li et al., 2022).

2.5.2 Alur Kerja PaddleOCR

Proses kerja PaddleOCR dimulai dengan memasukkan citra dokumen sebagai data masukan. Citra tersebut kemudian diproses oleh modul text detection untuk menemukan lokasi teks yang terdapat pada citra. Setelah area teks terdeteksi, sistem dapat melakukan direction classification untuk menentukan atau menyesuaikan orientasi teks. Area teks yang telah diproses kemudian diteruskan ke modul text recognition untuk mengenali karakter dan menyusunnya menjadi teks digital (Du et al., 2021).

Secara sederhana, alur kerja PaddleOCR dapat digambarkan sebagai berikut:

1. Citra Dokumen

   Merupakan citra masukan yang akan diproses oleh PaddleOCR, seperti foto atau hasil pemindaian dokumen ijazah.

2. Deteksi Teks

   Tahap ini digunakan untuk menemukan lokasi teks pada citra dan menentukan area yang mengandung informasi teks.

3. Klasifikasi Arah (*Direction Classification*)

   Tahap ini digunakan untuk menentukan orientasi teks yang telah terdeteksi sehingga teks dapat diproses sesuai dengan arah baca yang tepat.

4. Pengenalan Teks (*Text Recognition*)

   Area teks yang telah terdeteksi kemudian diproses untuk mengenali karakter dan menyusunnya menjadi kata atau kalimat.

5. Teks Digital

   Hasil akhir proses PaddleOCR berupa teks digital yang dapat ditampilkan, disimpan, dan digunakan untuk proses evaluasi dalam penelitian.

2.6 *Ground Truth*

Ground truth merupakan teks referensi yang dianggap benar dan digunakan sebagai acuan dalam mengevaluasi hasil pengenalan teks oleh sistem OCR. Dalam evaluasi OCR, hasil teks yang dihasilkan oleh sistem dibandingkan dengan *ground truth* untuk mengetahui tingkat kesalahan pengenalan pada tingkat karakter maupun kata. Oleh karena itu, *ground truth* menjadi bagian penting dalam proses evaluasi kinerja sistem OCR, khususnya ketika menggunakan metrik Character Error Rate (CER) dan Word Error Rate *(WER)* (Ströbel et al., 2020).

Pada penelitian ini, *ground truth* dibuat berdasarkan teks yang terdapat pada dokumen ijazah dan diverifikasi secara manual oleh peneliti. Proses pembuatannya dilakukan dengan membaca teks yang tercantum pada setiap dokumen ijazah, kemudian mengetik ulang informasi tersebut secara tepat ke dalam dokumen teks. Teks hasil pengetikan tersebut digunakan sebagai teks referensi yang mewakili informasi sebenarnya pada dokumen ijazah. Dengan

demikian, *ground truth* bukan merupakan hasil keluaran EasyOCR maupun PaddleOCR, melainkan teks referensi yang dibuat dan diperiksa berdasarkan dokumen sumber.

Setiap citra dokumen ijazah dalam dataset memiliki *ground truth* yang sesuai. Ketika citra yang sama diproses menggunakan EasyOCR dan PaddleOCR, hasil pengenalan teks dari kedua sistem OCR kemudian dibandingkan dengan *ground truth* untuk mengetahui kesalahan yang terdapat pada hasil pengenalan. Perbandingan antara hasil OCR dan teks referensi dapat digunakan untuk mengukur kualitas hasil OCR serta mengetahui tingkat kesalahan pengenalan teks (de Oliveira et al., 2023). Evaluasi dilakukan secara objektif karena kedua sistem diuji menggunakan acuan teks yang identik. Nilai CER dan WER kemudian digunakan untuk mengukur jumlah kesalahan pengenalan yang dihasilkan oleh masing-masing sistem. Nilai kesalahan yang lebih rendah menunjukkan bahwa hasil pengenalan teks lebih mendekati teks referensi.

2.7 *Character Error Rate (CER) dan Word Error Rate (WER)*

Character Error Rate (CER) dan *Word Error Rate* (WER) merupakan metrik yang digunakan untuk mengevaluasi tingkat kesalahan hasil pengenalan teks pada sistem *Optical Character Recognition* (OCR). Kedua metrik tersebut mengevaluasi hasil teks yang dihasilkan oleh sistem OCR dengan teks referensi atau *ground truth*. CER digunakan untuk mengukur kesalahan pada tingkat karakter, sedangkan WER digunakan untuk mengukur kesalahan pada tingkat kata. Penggunaan CER dan WER dapat memberikan gambaran mengenai tingkat

kesalahan pengenalan teks yang dihasilkan oleh sistem OCR (Sukarna et al., 2025; Ströbel et al., 2020).

2.7.1 Character Error Rate (CER)

Character Error Rate (CER) merupakan metrik yang digunakan untuk mengukur tingkat kesalahan pengenalan teks pada tingkat karakter. Perhitungan CER didasarkan pada tiga jenis kesalahan, yaitu substitusi (*substitution*), insersi (*insertion*), dan delesi (*deletion*). Substitusi terjadi ketika suatu karakter dikenali sebagai karakter lain, insersi terjadi ketika sistem menghasilkan karakter tambahan yang tidak terdapat pada *ground truth*, sedangkan delesi terjadi ketika karakter yang terdapat pada *ground truth* tidak berhasil dikenali oleh sistem OCR (Sukarna et al., 2025).

Rumus perhitungan CER adalah sebagai berikut:

$$CER = \frac{S + D + I}{N} \times 100\%$$

Keterangan:

- **S** = jumlah kesalahan substitusi;
- **D** = jumlah kesalahan delesi;
- **I** = jumlah kesalahan insersi;
- **N** = jumlah karakter pada *ground truth*.

Nilai CER menunjukkan tingkat kesalahan sistem OCR dalam mengenali karakter. Semakin rendah nilai CER, semakin sedikit kesalahan karakter yang dihasilkan sehingga hasil pengenalan teks semakin mendekati *ground truth*. Dalam penelitian ini, nilai CER digunakan untuk membandingkan tingkat

kesalahan karakter yang dihasilkan oleh EasyOCR dan PaddleOCR pada dokumen ijazah.

2.7.2 Word Error Rate (WER)

Word Error Rate (WER) merupakan metrik yang digunakan untuk mengukur tingkat kesalahan hasil pengenalan OCR pada tingkat kata. Berbeda dengan CER yang menghitung kesalahan pada setiap karakter, WER menghitung kesalahan berdasarkan kata. Kesalahan pada WER terdiri atas substitusi (*substitution*), insersi (*insertion*), dan delesi (*deletion*) terhadap kata yang terdapat pada *ground truth* (Sukarna et al., 2025).

Rumus perhitungan WER adalah sebagai berikut:

$$WER = \frac{S + D + I}{N} \times 100\%$$

Keterangan:

- **S** = jumlah kata yang salah dikenali atau digantikan;
- **D** = jumlah kata yang tidak berhasil dikenali;
- **I** = jumlah kata tambahan yang tidak terdapat pada *ground truth*;
- **N** = jumlah kata pada *ground truth*.

Semakin rendah nilai WER, semakin sedikit kesalahan kata yang dihasilkan oleh sistem OCR. Oleh karena itu, WER digunakan dalam penelitian ini untuk mengetahui kemampuan EasyOCR dan PaddleOCR dalam mengenali susunan kata pada dokumen ijazah.

Dalam penelitian ini, nilai CER dan WER dihitung dengan mengevaluasi hasil pengenalan teks dari EasyOCR dan PaddleOCR terhadap *ground truth* yang sama. CER digunakan untuk mengukur kesalahan pada tingkat karakter, sedangkan

WER digunakan untuk mengukur kesalahan pada tingkat kata. Hasil dari kedua metrik tersebut kemudian digunakan sebagai dasar perbandingan kinerja kedua sistem OCR. Sistem dengan nilai CER dan WER yang lebih rendah menunjukkan tingkat kesalahan pengenalan teks yang lebih rendah berdasarkan metrik tersebut (Sukarna et al., 2025).