

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil implementasi, pengujian, dan analisis komparatif EasyOCR dan PaddleOCR dalam pengenalan teks pada 100 citra dokumen ijazah, diperoleh kesimpulan sebagai berikut.

1. Kinerja EasyOCR dan PaddleOCR dalam mengenali teks pada dokumen ijazah menunjukkan hasil yang berbeda pada setiap dokumen yang diuji. Kedua sistem mampu melakukan pengenalan teks menggunakan citra masukan dan *ground truth* yang sama, tetapi menghasilkan tingkat kesalahan dan waktu pemrosesan yang berbeda. PaddleOCR cenderung menghasilkan pengenalan teks yang lebih baik karena memiliki tingkat kesalahan karakter dan kata yang lebih rendah serta waktu pemrosesan yang lebih singkat dibandingkan EasyOCR.
2. Berdasarkan perbandingan nilai CER, WER, dan waktu pemrosesan, PaddleOCR menunjukkan kinerja yang lebih baik dibandingkan EasyOCR. EasyOCR memperoleh rata-rata CER sebesar 0,2435 dan WER sebesar 0,4825, sedangkan PaddleOCR memperoleh rata-rata CER sebesar 0,2414 dan WER sebesar 0,4498. PaddleOCR menghasilkan CER lebih rendah pada 56 dokumen dan WER lebih rendah pada 58 dokumen. Pada aspek waktu pemrosesan, EasyOCR membutuhkan rata-rata 18,285 detik per dokumen, sedangkan PaddleOCR membutuhkan 1,240 detik per dokumen atau sekitar 14,75 kali lebih cepat. Dengan demikian, PaddleOCR lebih sesuai untuk

pengenalan teks pada dokumen ijazah berdasarkan dataset dan lingkungan pengujian yang digunakan dalam penelitian ini.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, beberapa saran yang dapat digunakan untuk pengembangan penelitian selanjutnya adalah sebagai berikut.

1. Penelitian selanjutnya dapat menggunakan jumlah dataset yang lebih besar dan lebih beragam, baik dari segi desain ijazah, jenis huruf, tata letak informasi, resolusi citra, pencahayaan, kemiringan dokumen, maupun kondisi pengambilan gambar. Hal tersebut dapat digunakan untuk memperoleh gambaran kinerja sistem OCR pada kondisi dokumen yang lebih bervariasi.
2. Tahap pra-pemrosesan citra dapat dikembangkan dengan menerapkan metode tambahan, seperti peningkatan kontras, pengurangan noise, koreksi kemiringan citra, atau peningkatan ketajaman citra. Pengujian terhadap beberapa metode pra-pemrosesan dapat dilakukan untuk mengetahui pengaruhnya terhadap nilai CER dan WER dari masing-masing sistem OCR.
3. Penelitian selanjutnya dapat membandingkan EasyOCR dan PaddleOCR dengan sistem OCR lainnya sehingga diperoleh perbandingan yang lebih luas mengenai kemampuan berbagai sistem OCR dalam mengenali teks pada dokumen ijazah.
4. Pengujian waktu pemrosesan dapat dilakukan pada beberapa spesifikasi perangkat keras dan lingkungan komputasi yang berbeda. Selain itu, pengujian masing-masing sistem OCR secara terpisah dapat dilakukan untuk mengetahui

kestabilan waktu pemrosesan tanpa adanya kemungkinan penggunaan sumber daya komputasi secara bersamaan.

5. Penelitian selanjutnya dapat mengembangkan aturan penyusunan ground truth dan normalisasi teks yang lebih terstandar, khususnya terkait penggunaan spasi, pergantian baris, huruf kapital, dan tanda baca. Hal tersebut dapat digunakan untuk mengetahui pengaruh perbedaan format teks terhadap hasil perhitungan CER dan WER.
6. Analisis hasil dapat dikembangkan menggunakan pengujian statistik lebih lanjut untuk mengetahui apakah perbedaan kinerja EasyOCR dan PaddleOCR bersifat signifikan secara statistik. Pengujian tersebut dapat melengkapi analisis statistik deskriptif yang digunakan dalam penelitian ini, terutama karena perbedaan rata-rata CER kedua sistem relatif kecil.
7. Sistem yang telah dikembangkan dapat ditingkatkan menjadi aplikasi yang tidak hanya melakukan pengenalan teks, tetapi juga mampu mengidentifikasi dan mengekstraksi informasi tertentu pada dokumen ijazah, seperti nama, nomor ijazah, tempat dan tanggal lahir, nama institusi, serta informasi akademik lainnya secara otomatis.