

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Bahasa Arab merupakan salah satu bahasa tertua di dunia, kaya akan sejarah dan pengetahuan. Banyak arkeolog yang masih bekerja keras untuk mengungkap fakta-fakta menariknya, mengingat akarnya yang dalam yang dapat ditelusuri kembali ke abad keenam. Sementara beberapa orang, khususnya mereka yang beragama Islam, mungkin telah mengenal bahasa Arab sejak kecil, orang lain yang baru mengenal bahasa tersebut mungkin merasa sistem penulisannya agak asing.

Tulisan tangan sulit untuk disimpan dan diakses secara efisien dan sesuai. Proses mengidentifikasi dan mengenali berbagai cara pada input seperti gambar, ucapan atau aliran karakter disebut *pattern recognition* (PR). Tahapan dalam proses tersebut dapat melibatkan pengukuran objek untuk mengidentifikasi atribut yang membedakan, ekstraksi fitur untuk atribut yang menentukan dan perbandingan dengan pola yang diketahui untuk menentukan apakah ada kecocokan antara keduanya atau tidak (ketidakcocokan).

Pengenalan karakter tulisan tangan Arab tetap menjadi masalah yang sulit karena kekhususan bahasa Arab. Alfabet Arab terdiri dari 28 huruf tanpa huruf besar dan kecil. Setiap huruf memiliki dua atau empat bentuk, tergantung pada posisinya dalam kata, yang mungkin berada di awal, tengah, akhir, dan terisolasi. Dalam alfabet Arab, 15 huruf dari 28 huruf memiliki satu titik atau lebih. Beberapa bentuk karakter serupa, tetapi perbedaan muncul dengan posisi dan jumlah titik yang berada di atas atau di bawah karakter menurut penelitian sebelumnya dengan judul “Investigation on Deep Learning for Off-line Handwritten Arabic Character Recognition” [1].

Meskipun sejumlah bahasa telah menerima perhatian yang signifikan dalam penelitian OCR. *Optical Character Recognition* (OCR) adalah teknologi yang dapat mengubah tulisan yang dicetak, ditulis tangan, atau diketik menjadi teks digital yang dapat dibaca mesin. Bahasa Arab masih sebagian besar merupakan bidang studi yang belum banyak dieksplorasi, dengan sebagian besar pekerjaan

sebelumnya pada bahasa ini berbasis karakter cetak. Pada penelitian “*Arabic Handwritten Characters Recognition using Convolutional Neural Network*” menggunakan metode *Convolutional Neural Network* (CNN) [2]. *Convolutional Neural Network* (CNN) adalah jenis khusus dari *multilayer feed-forward* yang dilatih dalam mode terawasi. CNN dilatih dan diuji menggunakan basis data yang berisi 16.800 karakter tulisan tangan Arab. Dalam makalah tersebut, metode optimasi diimplementasikan untuk meningkatkan kinerja CNN. Metode pembelajaran mesin umum biasanya menggunakan kombinasi antara ekstraktor fitur dan pengklasifikasi yang dapat dilatih. Penggunaan CNN menghasilkan peningkatan signifikan di berbagai pendekatan pembelajaran mesin.

Metode lain yang digunakan adalah penggunaan teknik *congealing* untuk memodelkan distorsi nyata pada tulisan tangan dan kemudian menggunakan model distorsi ini untuk mensintesis contoh tulisan tangan [3]. Adapun *Discrete Cosine Transform* (DCT) adalah teknik untuk mengonversi nilai *pixel* gambar menjadi komponen frekuensi dasarnya [4]. Dengan menerapkan DCT pada setiap gambar karakter, koefisien DCT dari gambar tersebut dapat diperoleh. Koefisien bernilai tinggi terkumpul di sudut kiri atas, sementara koefisien bernilai rendah berada di sudut kanan bawah array dan *Discrete Wavelet Transform* (DWT) adalah teknik lain yang digunakan untuk mengekstraksi fitur karakter, di mana pada setiap tingkat dekomposisi *filter low-pass* (LPF) dan *filter high-pass* (HPF) diterapkan pada setiap baris/kolom gambar untuk mendekomposisi gambar menjadi satu sub-pita frekuensi rendah (LL) dan tiga sub-pita frekuensi tinggi (LH, HL, HH) [5].

Dalam penelitian yang berjudul “Multi-Font Farsi/Arabic Isolated Character Recognition Using Chain Codes”, menerapkan menggunakan *chain code-based*. Diawali dengan menipiskan karakter tulisan, fitur vektor yang terdiri dari berbagai atribut dari setiap karakter Farsi/Arab diekstraksi. Bagian paling efisien dari fitur vektor karakter ini adalah *chain code* baru, yang dihitung menggunakan *neighborhood function* dan serangkaian normalisasi. Hasil eksperimen yang didapatkan menggunakan beberapa font Farsi standar menunjukkan bahwa metode yang diusulkan memiliki akurasi rata-rata 97,4% dalam pengenalan karakter yang benar [6].

Penelitian ini mengusulkan pendekatan OCR pada tulis tangan bahasa Arab menggunakan YOLO. YOLO adalah singkatan dari istilah 'You Only Look Once'. ini sering dikaitkan dengan aturan yang mendeteksi dan mengenali banyak objek dalam sebuah gambar (secara *real-time*). YOLOv9 disesuaikan dan dilatih untuk mengenali bahasa Arab. Namun, YOLO tidak menghasilkan tingkat pengenalan karena satu karakter mungkin memiliki beberapa hasil pengenalan. Penelitian ini meningkatkan OCR YOLOv9 dengan menggunakan tiga teknik: Menghapus hasil pengenalan yang berlebihan dengan memproses kotak yang tumpang tindih tumpang tindih dan menjaga kotak dengan akurasi tertinggi akurasi tertinggi. Selain itu, pemeriksa ejaan dengan jarak edit digunakan untuk mengoreksi kata-kata yang salah. Pada penelitian sebelumnya yang berjudul “Comprehensive Performance Evaluation of YOLOv11, YOLOv10, YOLOv9 and YOLOv8 on Detecting and Counting Fruitlet in Complex Orchard Environment” [7] membahas perbandingan performansi dari YOLO versi 8 sampai dengan versi 11, dimana YOLOv9 *Gelan-base* and YOLOv9 *Gelan-e* memiliki akurasi konfigurasi tertinggi dan YOLOv11 memiliki keunggulan dalam validasi penghitungan.

1.2. Rumusan Masalah

1. Dataset tulisan Arab yang belum representatif dan lengkap
2. Perlu adanya model yang efektif dalam deteksi tulisan Arab
3. Perlu adanya evaluasi kinerja terhadap model yang diusulkan dalam deteksi tulisan Arab

1.3. Tujuan dan Manfaat

1. Membangun data set yang lengkap dan representative dalam model deteksi tulisan tangan Arab.
2. Merancang model YOLOv9 dalam mendeteksi karakter Arab dengan tulisan tangan.
3. Membandingkan kinerja model yang diusulkan dengan model deteksi tulisan tangan lain yang telah ada.

1.4. Batasan Masalah

Ruang lingkup pembahasan dalam tugas akhir ini adalah sebagai berikut:

1. Aplikasi ini hanya dikembangkan untuk perangkat *Windows*.
2. Aplikasi pendeteksi difokuskan untuk huruf Hijaiyah botak tanpa menggunakan harakat.
3. OCR yang digunakan yaitu YOLOv9.
4. Deteksi dilakukan hanya pada tulisan tangan, bukan tulisan cetak.

1.5. Metode Penelitian

Metode penelitian yang diterapkan dalam tugas akhir ini, diantaranya sebagai berikut:

1. Studi Literatur

Tahap ini melakukan studi literatur dengan mengumpulkan referensi yang dibutuhkan untuk memperoleh informasi dan data yang berkaitan khususnya mengenai YOLOv9. Referensi dapat berupa buku-buku, artikel, maupun hasil penelitian yang berkaitan dengan penelitian.

2. Perancangan aplikasi YOLOv9.
3. Simulasi aplikasi.
4. Analisis dan evaluasi.

Tahap ini melakukan analisis performansi deteksi tulisan tangan yang dihubungkan dengan deteksi tulisan tangan YOLOv8 pada aplikasi.

5. Penarikan Kesimpulan

Tahap ini melakukan penarikan kesimpulan berdasarkan hasil simulasi dan analisis aplikasi yang telah dibuatkan. Kesimpulan dari tahap ini akan ditulis dalam laporan Tugas Akhir.