

Abstrak

Optical Character Recognition (OCR) merupakan salah satu aplikasi teknologi pada bidang pengenalan pola dan kecerdasan buatan sebagai mesin pembaca. Penelitian ini diajukan untuk meningkatkan kemampuan dalam mengenali huruf latin komputer yang merupakan standar huruf dokumen yang dicetak. Penelitian masih terus dilakukan untuk meningkatkan kemampuan sistem OCR yang masih terdapat beberapa kesulitan dalam sistem OCR. Seperti, keambiguan dalam teks dan *noise* dan kesalahan klasifikasi.

Untuk mengatasi permasalahan tersebut, penelitian ini mengajukan metode untuk meminimalkan kesalahan dengan mengoptimalkan ekstraksi ciri dengan vertex chain code (VCC) dan menggunakan Learning Vector Quantization (LVQ) sebagai metode klasifikasinya. Metode VCC mendapatkan ciri berdasarkan pada elemen rantai yang menunjukkan nilai asli yaitu mengindikasikan nomor sel vertex. LVQ melakukan pembelajaran pada lapisan kompetitif yang terawasi yang akan secara otomatis belajar untuk mengklasifikasikan vektor masukan dan selanjutnya kelas suatu karakter tergantung pada jarak antara vektor masukan tersebut.

Sistem OCR dengan metode VCC dan LVQ dapat mengenali huruf latin komputer dengan baik. Beberapa kesalahan klasifikasi terjadi karena huruf latin tersebut memiliki kemiripan bentuk atau ciri dan pemilihan ciri yang tidak cukup baik untuk membedakan kelas yang berbeda. Selain itu, sistem OCR yang dibangun ini belum dapat menangani keanekaragaman font, sehingga perlu dikembangkan klasifikasi yang dapat menangani lebih khusus untuk setiap kategori huruf latin.

Kata Kunci: OCR, Pengenalan Pola, Huruf Latin, *Vertex Chain Code*, LVQ