

Abstrak

Penelitian ini berfokus pada proses segmentasi citra dokumen teks tulisan tangan Aksara Jawa menggunakan algoritma Seglines dan dokumen yang digunakan adalah dokumen Serat Pertanda. Tujuan dilakukannya penelitian ini adalah untuk memahami konsep dasar algoritma *Seglines* pada citra dokumen tulisan tangan aksara Jawa dan menguji unjuk kerjanya.

Algoritma Seglines menghasilkan prosentase akurasi segmentasi sampai 96.45% - 99.79%. Prosentase itu didapat dari hasil segmentasi pada skrip yang digunakan secara luas di India, misalnya Gurmukhi. Algoritma ini dipilih karena dapat melakukan segmentasi terhadap baris yang bersentuhan secara horizontal dan menyatukan komponen dari baris yang *over segmented*.

Penelitian dilakukan dengan melakukan proses segmentasi pada citra dokumen yang telah mengalami proses normalisasi kemiringan, binerisasi dan reduksi *noise*. Penelitian kemudian dilanjutkan dengan menganalisa citra dokumen dan citra baris yang dihasilkan dari proses segmentasi dengan menggunakan algoritma Seglines.

Hasil pengujian yang dilakukan terhadap citra dokumen serat_pertanda2.jpg dengan menggunakan algoritma Seglines diketahui bahwa waktu yang diperlukan untuk memproses citra tersebut adalah 0,57486 detik. Prosentase akurasi yang didapatkan adalah 43,33 % untuk segmentasi citra dokumen tulisan tangan aksara Jawa dengan algoritma Seglines. Untuk hasil segmentasi setelah ditambahkan saran berupa fungsi penelusuran potongan aksara dengan garis potong segmentasi dan fungsi pencarian headlines (45% dari headlines) untuk segmentasi aksara didapatkan hasil 88,68% memenuhi syarat segmentasi. Prosentase hasil segmentasi dengan tolak ukur penulisan aksara Jawa didapatkan nilai sebesar 59,10% Kecilnya nilai prosentase disebabkan aksara yang sudah ditemukan ada beberapa bagiannya yang terpotong. Berdasarkan tinjauan di atas, maka dapat disimpulkan bahwa algoritma Seglines relatif baik untuk melakukan segmentasi citra dokumen teks tulisan tangan.

Abstract

This study focuses on the process of image segmentation of text documents using “Aksara Jawa” handwriting. “Serat Pertanda” document used by Seglines algorithms as input. The purpose of this study was to understand the basic concepts of Seglines algorithms on handwritten document image using “Aksara Jawa” handwriting and test its performance.

Research carried out by performing the document image segmentation process which has been processed using skew detection, binerization and noise reduction. The study was followed by analyzing the document image and the result.

The results of tests performed on the document image “serat_pertanda2.jpg” using the algorithm Seglines known that it takes 0.57486 seconds. The line image of the handwritten document image use segmentation algorithm Seglines obtained 17 lines correspond to the number of lines from the document image serat_pertanda2.jpg. This shows that the algorithm Seglines good if used for segmentation lines. Percentage of accuracy obtained is 43.33% for handwritten document image segmentation use algorithm Seglines base on Javanese script.

For the segmentation results after adding pieces of advice in the form of alphabet search function by secant segmentation and search functions headlines (45% of the headlines) for syllable segmentation results obtained 88.68%. This percentage value indicates that the system can find the main characters in each line. Percentage of segmentation results with a benchmark or writing the script of Java semantics obtained a value of 59.10% The low value is caused porsentase characters that have been found there are some parts that are cut off. Based on the review above, it can be concluded that the relatively good Seglines algorithm to segment the image of handwritten text documents.