

RINGKASAN

ANALISIS VARIASI METODE PENGHITUNGAN JARAK PADA PERFORMA ALGORITMA K-NEAREST NEIGHBOR DALAM KLASIFIKASI DETEKSI PRA-KANKER SERVIKS

Muhammad Satria Adjie Nugraha

Kanker serviks pada saat ini menempati urutan kedua sebagai penyakit kanker pada sistem reproduksi wanita dan peringkat keempat sebagai penyakit paling mematikan bagi wanita di seluruh dunia. Salah satu pencegahan yang telah dilakukan adalah dengan mendeteksi secara dini dari kanker tersebut atau deteksi pra-kanker menggunakan kolposkopi. Mengombinasikan kolposkopi dan perlakuan asam asetat pada serviks membantu tenaga medis dalam melihat adanya indikasi kanker berupa lesi. Dengan bantuan pengolahan citra, diharapkan proses deteksi pra-kanker ini dapat dilakukan dengan lebih efisien. Setelah citra kolposkopi melalui pre-processing dan seleksi fitur, dilakukan klasifikasi untuk membedakan citra serviks yang sehat dengan yang terindikasi adanya lesi. Dalam pendeteksian, salah satu algoritma klasifikasi yang dapat digunakan adalah metode *K-Nearest Neighbor*. Penggunaan metode ini melibatkan perhitungan jarak dari data baru terhadap sejumlah (nilai K) data terdekat. Karena terdapat berbagai metode penghitungan jarak untuk *K-Nearest Neighbor*, perlu diteliti metode yang cocok dalam kasus klasifikasi ini.

Tahapan penelitian ini meliputi tahap persiapan alat dan bahan, serta studi literatur terkait. Berikutnya dilanjut dengan tahap *preprocessing* dan segmentasi untuk meminimalisasi derau. Setelah itu akan dilakukan proses ekstraksi fitur dan dilanjut dengan klasifikasi. Dalam klasifikasi menggunakan algoritma *K-Nearest Neighbor* metode penghitungan jarak yang akan digunakan meliputi Euclidean, Manhattan, Bray-Curtis, Canberra, Chebyshev, Cosine, Mahalanobis, dan Minkowski. Setelah citra melalui proses ekstraksi fitur, akan dilakukan uji coba dan evaluasi hasil klasifikasi pada model *K-Nearest Neighbor* dengan metode perhitungan jarak yang berbeda. Hasil akhir menunjukkan bahwa metode jarak Chebyshev merupakan metode yang paling optimal untuk aplikasi pada model klasifikasi *K-Nearest Neighbor* untuk mendeteksi pra-kanker serviks citra kolposkopi dengan tingkat akurasi 97%.

Kata kunci : Klasifikasi, *K-Nearest Neighbor*, pra-kanker serviks, penghitungan jarak

SUMMARY

PERFORMANCE ANALYSIS OF K-NEAREST NEIGHBOR CLASSIFIER ALGORITHM WITH VARYING DISTANCE MEASURES FOR PRE-CERVICAL CANCER DETECTION

Muhammad Satria Adjie Nugraha

At the moment, cervical cancer is ranked second in the most number of cancer cases in women's reproduction system and fourth as the most lethal disease for woman around the world. One of the preventive approach to this issue is by conducting a pre-cancer detection using colposcopy. Combining both colposcopy and the usage of acetic acid solution, medical personnel can see indications of pre-cancer in the form of lesions on the cervix. With the help of image processing, the efforts of pre-cervical cancer detection can hopefully be done in a more efficient manner. After colposcopy images went through pre-processing and feature extraction, classification will be done in order to differentiate between a healthy cervix image and lesion-affected cervix image. In the process of detection, one of the widely known algorithms used for classification is called K-Nearest Neighbor. This algorithm involves distance calculation of a new data relative to a certain number (the value of K) of the existing data. Since there are various measures of distance calculation, this study aims to find the most suitable distance measure for this certain application.

The distance measures used on this paper involves Euclidean, Manhattan, Bray-Curtis, Canberra, Chebyshev, Cosine, Mahalanobis, dan Minkowski. After the images went through feature extraction process, testing and evaluation of classification results for the K-Nearest Neighbor classifier with varying distance measures will be done. Final results show that the optimal distance measure for this specific application of pre-cervical cancer detection using K-Nearest Neighbor algorithm is the Chebyshev distance measure with 97% accuracy rate.

Keywords : Classification, K-Nearest Neighbor, pre-cervical cancer, distance measure