

ABSTRAK

Klasifikasi daun herbal berbasis citra dilakukan berdasarkan fitur tekstur *Local Binary Pattern* dan fitur bentuk *Hu Moment* kemudian diklasifikasikan menggunakan algoritma KNN. Fitur bentuk pada daun dapat direpresentasikan secara numerik dengan menggunakan *Hu Moment*, namun *Hu Moment* memiliki kelemahan ketika mengenali daun dari spesies berbeda yang memiliki bentuk hampir sama. Oleh karena itu, LBP digunakan untuk mengekstrak fitur tekstur pada daun. Dengan menggabungkan kedua fitur tersebut, akan menghasilkan fitur yang lebih beragam dan memungkinkan sistem untuk melakukan klasifikasi daun herbal dengan lebih akurat. Namun, penggabungan kedua fitur juga dapat meningkatkan kompleksitas data dan berpengaruh terhadap kinerja algoritma KNN. Untuk mengurangi kompleksitas tersebut, salah satu caranya yaitu dengan mengurangi dimensi vektor.

Seleksi fitur merupakan teknik yang dapat digunakan untuk mengurangi dimensi vektor dengan menghapus atribut yang kurang relevan. Salah satu metode seleksi fitur, yaitu *Information Gain*, merupakan metode yang bekerja dengan cara melakukan pemeringkatan fitur yang paling berpengaruh atau paling banyak memiliki informasi terhadap kelas tertentu. Penggunaan seleksi fitur *Information Gain* dapat menghasilkan fitur yang paling relevan dan mengurangi tingkat kompleksitas suatu algoritma klasifikasi serta menghasilkan interpretasi model yang lebih baik.

Pengujian dilakukan berdasarkan *confusion matrix* untuk menghitung nilai akurasi dari model klasifikasi daun herbal yang meliputi pengujian nilai k, kombinasi fitur, dan penerapan seleksi fitur. Hasil pengujian menunjukkan akurasi yang berbeda-beda pada setiap nilai k dan fitur yang digunakan. Pada pengujian tanpa menggunakan seleksi fitur, nilai akurasi tertinggi yang didapat adalah ketika mengkombinasikan kedua fitur LBP dan *Hu Moment* dengan nilai k=3, yaitu sebesar 89,58%. Sedangkan pada pengujian menggunakan seleksi fitur, nilai akurasi tertinggi yang didapat adalah ketika menggabungkan kedua fitur dan jumlah fitur yang digunakan sebanyak 14 fitur dengan nilai k=1, yaitu sebesar 90,83%. Namun, jika penggunaan seleksi fitur hanya pada salah satu fitur LBP atau *Hu Moment* saja, tidak meningkatkan akurasi bahkan dapat menghasilkan akurasi yang lebih rendah.

Kata kunci: Daun herbal, ekstraksi fitur, seleksi fitur, klasifikasi, *Local Binary Pattern*, *Invariant Hu Moments*, *Information Gain*, *K-Nearest Neighbor*

ABSTRACT

Image-based herbal leaf classification is carried out based on *Local Binary Pattern* texture features and *Hu Moment* shape features and then classified using the *K-Nearest Neighbor* algorithm. Shape features on leaves can be represented numerically using *Hu Moment*, but *Hu Moment* has a weakness when recognising leaves from different species that have almost the same shape. Therefore, *LBP* is used to extract the texture features of the leaves. By combining the two features, it will produce more diverse features and allow the system to perform herbal leaf classification more accurately. However, combining both features can also increase data complexity and affect the performance of the *KNN* algorithm. To reduce the complexity, one way is to reduce the dimension of the vector.

Feature selection is a technique that can be used to reduce vector dimension by removing irrelevant attributes. One of the feature selection methods, *Information Gain*, is a method that works by ranking the features that have the most influence or the most information on a particular class. The use of *Information Gain* feature selection can produce the most relevant features and reduce the complexity of a classification algorithm and produce better model interpretation.

Tests were conducted based on the *confusion matrix* to calculate the accuracy value of the herbal leaf classification model which includes testing the *k* value, feature combination, and application of feature selection. The test results show different accuracy for each *k* value and feature used. In testing without using feature selection, the highest accuracy value obtained is when combining both *LBP* and *Hu Moment* features with a value of $k = 3$, which is 89.58%. Meanwhile, if only using the *LBP* texture feature, the best accuracy produced is at $k = 1$ for 80%. Then if only the *Hu Moment* shape feature, the best accuracy obtained is only 70.83% when the value of $k = 11$. Then in testing using feature selection, the highest accuracy value obtained is when the number of features used is 14 features with a value of $k = 1$, which is 90.83%. Thus, the results show that by combining two feature extractions plus feature selection can produce better accuracy than using only one feature or using both features without feature selection.

Keyword: Herbal leaf, feature extraction, feature selection, classification, *Local Binary Pattern*, *Invariant Hu Moments*, *Information Gain*, *K-Nearest Neighbor*