Revolutionize Date Fruit Classification using Optimized K-Nearest Neighbor

Khaidir Mauladan[1], Wikky Fawwaz Al Maki[2]

[1,2]School of Computing, Telkom University, Bandung
[1]khaidir@students.telkomuniversity.ac.id, [2]wikkyfawwaz@telkomuniversity.ac.id

Abstrak
Berbagai varian buah kurma yang menyebar di seluruh dunia memiliki kompleksitas dan karakteristik yang unik seperti warna, rasa, bentuk, dan tekstur. Studi ini mempertimbangkan fitur-fitur yang rumit dan beragam dari berbagai spesies buah kurma. Ciri-ciri unik ini dapat menimbulkan tantangan dalam menentukan dan membedakan antara berbagai jenis buah kurma yang ada. Klasifikasi buah-buahan ini dapat menjadi sangat sulit karena perbedaan-perbedaan halus dalam fitur yang telah disebutkan yang ada di antara spesies yang berbeda. Untuk mengatasi masalah ini, kategorisasi otomatis buah kurma telah muncul sebagai hasil dari pengembangan pembelajaran mesin dan visi komputer. Studi ini mengusulkan skema klasifikasi lima kategori untuk buah kurma. Untuk menghasilkan data pelatihan berkualitas tinggi untuk pengklasifikasi K-Nearest Neighbor (KNN), ekstraksi fitur dilakukan pada gambar buah kurma menggunakan momen warna, circularity, dan deskriptor HOG masing-masing. Analisis Komponen Utama (PCA) digunakan sebagai pendekatan reduksi dimensionalitas untuk meningkatkan model KNN yang diusulkan. Untuk menentukan fitur terbaik yang ditentukan berdasarkan pengaruhnya pada kinerja model, Binary Particle Swarm Optimization (BPSO) digunakan sebagai pendekatan pemilihan fitur. Model klasifikasi mencapai tingkat akurasi 93,08%, yang merupakan peningkatan yang cukup besar sebesar 10,77% dibandingkan dengan model KNN dasar.

Kata Kunci: Binary Particle Swarm Optimization, Circularity, Color Moments, Histogram of Oriented Gradients, K-Nearest Neighbor, Principal Component Analysis

Abstract
Different variations of date fruit spreading around the globe possess their unique complexity and distinctive characteristics such as color, taste, shape, and texture. This study took into account the intricate and diverse features of various species of date fruit. These unique traits can present challenges in accurately defining and distinguishing between the numerous types of date fruit that exist. The classification of these fruits can be particularly difficult due to the subtle differences in the aforementioned features that exist between the different species. To solve this issue, automatic date fruit categorization has emerged as a result of the development of machine learning and computer vision. This study proposes a five-category classification scheme for date fruit. To generate high-quality training data for the K-Nearest Neighbor (KNN) classifier, feature extraction was done on the date fruit images using color moments, circularity, and HOG descriptors respectively. Principal Component Analysis (PCA) was employed as a dimensionality reduction approach to improving the suggested KNN model. To determine the best features determined by their influence on the model's performance, Binary Particle Swarm Optimization (BPSO) was used as a feature selection approach. The classification model attained an accuracy rate of 93.08%, which is a considerable increase of 10.77% over the basic KNN model.

Keywords: Binary Particle Swarm Optimization, Circularity, Color Moments, Histogram of Oriented Gradients, K-Nearest Neighbor, Principal Component Analysis

1. Introduction

Background

Date fruits are a good source of nourishment and are commonly grown in hot and dry areas across the world [22]. Accurate and dependable fruit categorization is critical for improving the efficiency of date fruit production and quality management. Date fruit has an extensive cultivation history stretching back to previous civilizations. Date fruit variations abound over the world, including but not limited to Galaxy, Sugaey, Shaishe, Ajwa, and Nabtat Ali, each with its own particular texture, flavor, color, and shape. However, there are various obstacles to overcome while classifying date fruits. One difficulty is the visual variety of date fruits, since different varieties