

PREDIKSI VOLUME SANTAN KELAPA MENGGUNAKAN ANDROID DENGAN METODE ANALISIS WARNA BERBASIS PENGOLAHAN CITRA DIGITAL

Annisa Budiastri¹, Achmad Rizal², Ratri Dwi Atmaja³

¹Teknik Telekomunikasi, Fakultas Teknik Elektro, Universitas Telkom

Abstrak

Kelapa dengan nama latin *Cocos Nucifera* adalah tumbuhan yang dimanfaatkan hampir semua bagiannya oleh manusia sehingga dianggap sebagai tumbuhan serba guna, kelapa juga adalah sebutan untuk buah yang dihasilkan dari tumbuhan ini. Bagian yang paling utama dimanfaatkan adalah daging buahnya sebagai bahan baku pembuatan santan. Untuk memperoleh santan yang memiliki volume yang banyak harus diperhatikan kematangan buah kelapa tersebut. Proses pemilihan buah kelapa bergantung kepada persepsi manusia terhadap komposisi warna kecoklatan kulit ari daging buah yang dimiliki kelapa. Namun proses identifikasi menurut persepsi manusia memiliki kelemahan diantaranya kelelahan juga perbedaan persepsi tiap manusia.

Tugas akhir ini dirancang sebuah sistem klasifikasi untuk mengetahui volume santan buah kelapa berdasarkan warna kecoklatan dari kulit ari buah dengan menggunakan metode analisis warna dan berbasis pengolahan citra digital. Data citra yang diambil adalah buah kelapa yang telah dikupas dari batoknya dengan menggunakan kamera ponsel android 8 MP.

Hasil dari pengolahan citra digital tersebut akan menjadi input pada pengenalan pola dan identifikasi buah kelapa, sehingga dapat diketahui karakteristik yang ada pada setiap buah kelapa. Tingkat akurasi aplikasi CoconutMilk secara keseluruhan adalah 80% dengan waktu komputasi 44,95 detik .

Kata Kunci : Analisis warna, kelapa, santan, pengolahan citra digital

Abstract

Coconut or the Latin name *Cocos Nucifera* is a plant which almost parts of it utilized by man so it is considered as a multipurpose plant, coconut is also the term for the fruit which produced from this plant. The most important part is the pulp used as the material for coconut milk. To obtain the volume of coconut milk we have to observe the maturity of the coconut. The process of selecting coconuts depends on a man perception about the composition of brown color of coconut epidermis. But the identification process according to man perception has disadvantages such as fatigue and also the differences perception of each human being.

This thesis designed a classification system to determine the volume of coconut milk based on the brown color of epidermis using a color analysis method and based on digital image processing. The image data taken from the coconut which was peeled from the coconut shell using 8 Megapixel camera of android phones.

The result of digital image processing will be input pattern recognition and identification of coconut, so the characterizing that exists in each coconuts can be known. The accuracy of CoconutMilk application is 80% overall with a time of 44,95 second computing.

Keywords : Color analysis, coconut, coconut milk, digital image processing

BAB I

PENDAHULUAN

1.1. Latar Belakang

Analisis diskriminan merupakan teknik menganalisis data, dimana variabel dependen merupakan data kategori (nominal dan ordinal) sedangkan variabel independen berupa data interval atau rasio. Analisis diskriminan bertujuan untuk mengklasifikasikan suatu individu atau observasi ke dalam kelompok yang saling bebas (*mutually exclusive*) atau dikenal juga dengan metode dependensi yaitu variabel-variabelnya tidak saling bergantung satu dengan yang lain. Berbeda dengan metode interdependensi yaitu metode yang antarvariabelnya ada saling ketergantungan. Sistem tersebut juga dapat menganalisa dan mengidentifikasi bentuk, tekstur, atau bahkan warna dari suatu objek.

Dalam industri telekomunikasi, penerapan analisis diskriminan pun sangat diperlukan. Berbagai macam data yang penting mengenai infrastruktur pembangunan jaringan telekomunikasi maupun data yang berkaitan dengan pelanggan akan dikumpulkan dalam suatu database yang dikumpulkan oleh manusia atau pekerja industri telekomunikasi tersebut. Hal ini kerap kali menimbulkan kesalahan atau *human error* secara tidak sengaja. Meskipun jumlah error sedikit, tetapi cukup berdampak pada infrastruktur telekomunikasi lain yang masih berkaitan. Maka analisis perlu dilakukan agar dapat diketahui letak data yang error atau bersifat sebagai *outlier*.

Sebagai suatu metode yang alternative dibandingkan metode-metode yang sudah ada digunakan metode yang memanfaatkan hubungan antara fitur-fitur pada suatu kelas untuk mengklasifikasikan suatu sampel pada kelas tertentu dengan data yang berjumlah besar yang tidak dapat dengan mudah diselesaikan oleh metode-metode klasifikasi yang lain. Maka akan dilakukan suatu analisis yang bertujuan untuk merealisasikan penggunaan metode Linear Discriminant Analysis dalam keakurasian klasifikasi data termasuk menentukan munculnya *data error* dalam bentuk outlier dengan menggunakan perangkat lunak MATLAB. Dari keunggulan metode Linear Discriminant Analysis ini kemudian akan diaplikasikan Cross Validation. Untuk membuktikan kemampuan klasifikasi, akan dilakukan pengujian pada metode ini dengan menggunakan data set *Wine*. Penggunaan dataset tersebut dikarenakan

inter-relation fitur pada data set tersebut lebih kuat daripada contoh data set lain. Hal ini semakin menunjang untuk menampilkan keunggulan penggunaan metode klasifikasi.

Pembangunan klasifikasi dataset dengan menggunakan metode Linear Discriminant Analysis dan *Principal Component Analysis* dengan *Cross Validation* ini dibagi menjadi dua bagian, yaitu tahap pelatihan dan tahap pengujian. Tahap pelatihan bertujuan untuk mendapatkan fitur-fitur penting yang akan menjadi masukan untuk MATLAB, dimana pada tahap pelatihan proses klasifikasi ini bertujuan untuk mendapatkan *classifier* yang dapat memisahkan kelas-kelas yang ada pada dataset. Kemudian tahap pengujian bertujuan untuk melihat bagaimana sistem berjalan dari awal sampai akhir dengan beberapa parameter pengujian. Diharapkan dengan adanya penelitian ini dapat menghasilkan sistem klasifikasi data dengan akurasi yang baik dan penggunaan metode yang tepat.

1.2. Tujuan

Adapun tujuan dari penelitian dan analisis dari Tugas Akhir ini adalah:

1. Melakukan pengklasifikasian data menggunakan Linear Discriminant Analysis yang bersifat distribusi normal dan homoskedastik yang dapat memaksimalkan sistem pengenalan pola dalam mengklasifikasikan data.
2. Mengklasifikasikan data dari dataset ke dalam bentuk yang ternormalisasi, baik pada Linear Discriminant Analysis maupun pada Principal Component Analysis.
3. Mengaplikasikan teknik klasifikasi yang telah ternormalisasi dalam metode *Cross Validation*.

1.3. Rumusan Masalah

Permasalahan yang dijadikan obyek penelitian dan pengembangan pada Tugas Akhir ini adalah:

1. Bagaimana mengklasifikasikan data dengan menggunakan metode Linear Discriminant Analysis dan Principal Component Analysis?
2. Bagaimana membentuk sebuah dataset ke dalam bentuk data yang telah ternormalisasi?

3. Bagaimana merancang algoritma metode klasifikasi pada perangkat lunak MATLAB?

1.4. Batasan Masalah

Untuk menghindari meluasnya materi dalam pembahasan Tugas Akhir ini, maka penulis membatasi permasalahan dalam Tugas Akhir ini dengan hanya mencakup hal-hal berikut:

- a. Data yang digunakan dalam penelitian berasal dari UCI Repository.
- b. Penelitian ini menggunakan metode Linear Discriminant Analysis.
- c. Klasifikasi yang telah dihasilkan akan dinormalisasi dan diiterasi dengan teknik Cross Validation.

1.5. Metode Penelitian

Metode penelitian yang digunakan untuk menyelesaikan tugas akhir ini adalah metode ekperimental, sedangkan prosedur penelitian untuk menyelesaikan Tugas Akhir ini adalah:

- a. Studi Literatur
Prosedur ini dilakukan untuk mempelajari konsep dan teori-teori yang mendukung proses perancangan sistem
- b. Survey lapangan
Analisis permasalahan dan kebutuhan
- c. Perancangan dan realisasi
Merancang suatu analisis dengan menggunakan algoritma pada *software* MATLAB
- d. Analisis dan pengujian
Berdasarkan teori yang sudah diperoleh sebelumnya, kemudian melakukan pengujian terhadap rancangan analisis pada *software*.
- e. Konsultasi
Dilakukan dengan dosen pembimbing menyangkut petunjuk dan pertimbangan-pertimbangan praktis mengenai perancangan analisis tersebut.

1.6. Sistematika Penulisan

Penulisan proposal ini akan dibagi menjadi beberapa bagian sebagai berikut:

BAB I Pendahuluan

Berisi latar belakang, tujuan, rumusan masalah, batasan masalah, metode penelitian, dan sistematika penulisan

BAB II Dasar Teori

Mengenai dasar-dasar teori yang diperlukan serta literature-literatur yang mendukung dalam penelitian dan analisis penggunaan metode Linear Discriminan Analysis pada sistem *error rate classifier*.

BAB III Perancangan dan Implementasi

Mengenai pembahasan perancangan awal system yang kemudian diimplementasikan dengan menggunakan algoritma pada perangkat lunak MATLAB.

BAB IV Pengujian dan Analisis Hasil Implementasi

Membahas mengenai pengujian dan analisis hasil implementasi yang telah dilakukan sebelumnya. Hal ini bertujuan untuk mengetahui kecocokan antara analisis yang dilakukan secara teori dengan menggunakan perangkat lunak MATLAB.

BAB V Kesimpulan dan Saran

Berisi kesimpulan dari penelitian Tugas Akhir ini serta saran yang diperlukan untuk pengembangan sistem lebih lanjut.

BAB V

KESIMPULAN DAN PENUTUP

5.1 Kesimpulan

Berdasarkan hasil pengujian dan analisis yang telah dilakukan, dapat disimpulkan beberapa hal berikut ini:

1. Penerapan *cross validation* pada metode klasifikasi *Linear Discriminant Analysis* berhasil dilakukan. Dan dari data pengujian yang diperoleh, hasil penerapan *cross validation* pada teknik klasifikasi meningkatkan tingkat akurasi hasil klasifikasi dibandingkan tanpa menerapkan teknik *cross validation* pada *Linear Discriminant Analysis*, begitu pula pada *Principal Component Analysis*.
2. Tingkat akurasi penyebaran dengan metode PCA dapat dikatakan tidak baik, dikarenakan sifat distribusi PCA dimana rata-rata dan kovariansinya tidak resisten pada distribusi data. Seperti pada Tabel 4.5 dapat diketahui bahwa nilai terkecil **1.9** yang diperoleh LDA dengan normalisasi menunjukkan bahwa metode LDA lebih baik dibanding PCA yaitu sebesar **5.9**.
3. Sebelum dilakukan Cross validation dan klasifikasi, harus dilakukan proses normalisasi data terlebih dulu pada dataset agar persebaran data merata dengan menggunakan algoritma *zscore*, dimana cakupan nilai rata-rata penyebaran antara 0 hingga 1.
4. Penerapan Cross validation menggunakan toolbox *crossvalind* pada data yang sudah distribusi secara LDA.

5.2 Saran

Berikut ini merupakan beberapa hal yang disarankan penulis untuk dilakukan pada penelitian yang sehubungan dengan penelitian ini di masa mendatang:

1. Dengan penerapan Cross validasi ini, penelitian dapat dilanjutkan pada data yang bersifat non-linear.

2. Untuk menambah tingkat akurasi sistem pada klasifikasi data, sebaiknya penelitian dilakukan pada jenis dataset yang lain agar hasil yang diperoleh pun lebih banyak dan beragam, serta menambah tingkat kepastian dalam penggunaan teknik *cross validation*.
3. Untuk penelitian lebih lanjut dapat digunakan metode lain selain *Linear Discriminant Analysis* maupun *Principal Component Anlysis* sebagai pembandingan terhadap hasil yang telah diperoleh dari penelitian saat ini.



DAFTAR PUSTAKA

- [1] Hart, Duda. *Pattern Classification*. Diperoleh 28 Januari 2013, dari <http://syridink.com/files/159Download%20pattern%20classification%20.pdf>
- [2] Haihong ZHANG, Cuntai GUAN, and Kai Keng Ang. 2010, *An Information Theoretic Linear Discriminant Analysis Method*. Diperoleh 29 Januari 2013, dari <http://www.ieeexplore.ieee.org:80?xpl/articleDetails.jsp?reload=true&arnumber=5597750>.
- [3] *Analisis Diskriminan*. Diperoleh pada 3 Februari 2013. Dari http://daps.bps.go.id/file_artikel/65/ANALISIS%20DISKRIMINAN.pdf
- [5] Godam64. (2010, 21 Juni), *Klasifikasi Jenis, dan Macam Data – Pembagian Data Dalam Ilmu Eksak*. Diperoleh 13 Maret 2013, dari http://organisasi.org/klasifikasi_jenis_dan_macam_data_pembagian_data_dalam_ilmu_eksak_sains_statistik_statistika
- [7] *Machine Learnig Repositroy: Wine Data Set*. Diperoleh pada 14 Maret 2013. Diperoleh dari <http://archive.ics.uci.edu/ml/datasets/Wine>
- [8] *Validation*. Diperoleh pada 15 April 2013. Diperoleh dari http://research.cs.tamu.edu/prism/lectures/iss/iss_113.pdf
- [9] Natalius, Samuel. *Metoda Naive Bayes Classifier dan Penggunaannya pada Klasifikasi Dokumen*, Makalah II2092 Probabilitas dan Statistik – Sem. I Tahun 2010/2011, Bandung. Diperoleh pada 16 April 2013.

- [10] Scholdz, Matthias. *Principal Component Analysis*. Diperoleh pada 29 April 2013. Dari http://www.nlpca.org/pca_principal_component_analysis.html
- [12] *Principal Component Analysis and Linear Discriminant Analysis with GNU Octave*, Diperoleh 27 Mei 2013 dari http://www.bytefish.de/blog/pca_lda_with_gnu_octave/
- [13] Shlenz, Jonathan. *Principal Component in Matla*. Diperoleh pada 4 Juni 2013. Dari http://www.cs.stevens.edu/~mordohai/classes/cs559_s09/PCA_in_MATLAB.pdf
- [14] *DataSet Object Examples*. Diperoleh pada 4 Juni 2013. Dari wiki.eigenvector.com/index.php?title=DataSet_Object_Examples
- [15] *Principal Component Analysis of Raw Data*. Diperoleh pada 10 juni 2013. Dari <http://www.mathworks.com/help/stats/pca.html#bti6n7k-2>
- [16] *Linear Discriminan Analysis for Dataset*. Diperoleh pada 10 Juni 2013. Dari <http://www.mathworks.com/help/stats/lda-for-dataset.html#ok4f7n-8>
- [17] *Cross Validation*. Diperoleh pada 8 September 2013. Dari http://www.cse.iitb.ac.in/~tarung/smt/papers_ppt/ency-cross-validation.pdf

Telkom
University