

Penggunaan Principal Component Analysis dan Minimum Distance Classifier untuk Pengenalan Citra Buah

Bella Adinda Putri¹
Fakultas Ilmu Komputer
Universitas Sriwijaya
Palembang, Indonesia
belladinda92@gmail.com

Yulia Hapsari²
Fakultas Ilmu Komputer
Universitas Sriwijaya
Palembang, Indonesia
yuliahapsari.str@gmail.com

Abstract—Pengenalan buah dengan alat yang dapat diandalkan merupakan sebuah tantangan, walaupun manusia bisa mengenali buah-buahan hampir dengan tanpa usaha. Untuk mencapai tujuan tersebut, perlu dilakukan studi kepustakaan untuk memahami konsep dan landasan teori agar dapat memperkuat asumsi metode *Principal Component Analysis* dan *Minimum Distance Classifier*. Fokus utama dalam penelitian ini adalah bagaimana mendapatkan fitur dari setiap citra buah untuk membedakan buah satu sama lain dengan menerapkan metode *Principal Component Analysis* (PCA) sebagai ekstraksi ciri, dan *Minimum Distance Classifier* sebagai algoritma pengenalan sehingga didapatkan hasil pengenalan yang akurat. Hasil pengenalan dengan menggunakan data latih (*training data set*) mendapatkan keakuratan sebesar 100%, sedangkan hasil pengenalan menggunakan data uji (*testing data set*) mendapatkan keakuratan sebesar 84%. Sehingga dapat disimpulkan bahwa ekstraksi ciri menggunakan PCA dapat digunakan dalam penerapan algoritma *Minimum Distance Classifier* untuk pengenalan buah.

Keywords—*pengenalan buah; ekstraksi ciri; principal component analysis; minimum distance classifier*

I. PENDAHULUAN

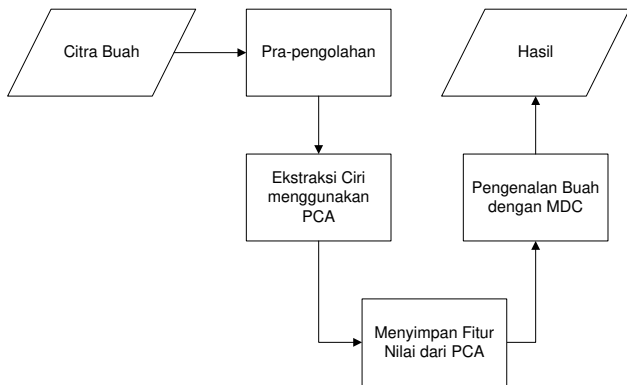
Pengenalan pola adalah disiplin ilmu yang tujuannya untuk mengklasifikasi objek atau pola menjadi beberapa kategori atau kelas [1]. Dari pengenalan suara otomatis, identifikasi sidik jari, pengenalan karakter optik, identifikasi urutan DNA, dan lain sebagainya, dapat terlihat jelas bahwa akurasi pengenalan pola dengan mesin akan sangat berguna. Selain itu, dalam memecahkan masalah-masalah yang berbeda yang dibutuhkan untuk membangun sistem tersebut, kita

mendapatkan pemahaman yang lebih dalam dan apresiasi untuk sistem pengenalan pola. Untuk beberapa masalah, seperti pengenalan suara dan pengenalan visual, upaya desain dapat dipengaruhi oleh pengetahuan tentang bagaimana hal ini diselesaikan secara alami, baik dalam algoritma dan dalam perangkat keras yang dirancang khusus [2].

Sistem pengenalan telah muncul sebagai 'tantangan besar' dalam visi komputer, dengan tujuan jangka panjang untuk bisa mencapai dan mendekati tingkat pengenalan oleh manusia untuk puluhan ribu kategori di bawah berbagai kondisi [3]. Tantangan utama adalah mencapai *pre-processing* optimal, fitur ekstraksi pilihan, dan klasifikasi yang benar, terutama dalam kondisi variabilitas input data [4].

Sebuah sistem pengenalan buah adalah sistem berbasis komputer yang mampu secara otomatis mengidentifikasi citra buah yang berbeda. Sistem ini bergantung pada algoritma pengenalan, seperti citra buah Eigen [5]. Sistem pengenalan buah dapat diterapkan untuk tujuan pendidikan dalam meningkatkan pembelajaran, terutama untuk anak-anak kecil dan pasien keterbelakangan mental, berdasarkan hasil pengenalan buah [6]. Dalam penelitian ini, sistem pengenalan pola yang akan dibangun adalah pengenalan citra buah menggunakan *Principal Component Analysis* sebagai ekstraksi ciri dan *Minimum Distance Classifier* sebagai metode pengenalan.

Pokok perhatiannya adalah bagaimana mendapatkan karakteristik gambar yang dapat membedakan antara buah satu dengan yang lainnya. Hasilnya merupakan sistem berbasis komputer yang dapat digunakan untuk mengenali buah dengan menggunakan citra buah itu sendiri. Gambar 1 merupakan mekanisme sistem pengenalan buah.



Gambar 1 Mekanisme Sistem Pengenalan Buah

II. METODE PENELITIAN

2.2. Principal Component Analysis

Metode ekstraksi ciri yang digunakan dalam pengenalan citra buah ini adalah *Principal Component Analysis* (PCA). Tahap pertama dalam perhitungan PCA, setelah citra buah telah melalui proses *cropping*, nilai piksel citra dari kiri atas sampai kanan bawah diambil dan diubah ke dalam bentuk vektor.

Tahap kedua adalah menghitung besar rata-rata dari nilai vektor masing-masing gambar. Proses ini dilakukan dengan menjumlahkan masing-masing vektor pada setiap citra buah dan dibagi dengan panjang vektor untuk didapatkan nilai rata-rata yang diinginkan. Proses ini dapat dilakukan dengan menggunakan persamaan

$$\mu = \frac{1}{M} \sum_{i=1}^M x_i \quad (1)$$

dan untuk pengurangan nilai vektor ke -n dengan rata-ratanya digunakan persamaan

$$\varphi_i = x_i - \mu \quad (2)$$

Tahap ketiga yaitu proses perhitungan matrik kovarian, hasil yang telah didapatkan dari pengurangan nilai vektor ke -n dengan rata-ratanya dikalikan dengan *transpose* nya dengan menggunakan persamaan

$$C = \frac{1}{n-1} \sum_{i=1}^M (x_i - \mu) (x_i - \mu)^T \quad (3)$$

sehingga akan didapatkan sebuah matriks yang disebut dengan matriks kovarian. Tahap keempat adalah menghitung *Eigenvalue* menggunakan persamaan

$$\text{Det} (\chi \cdot I - C) = 0 \quad (4)$$

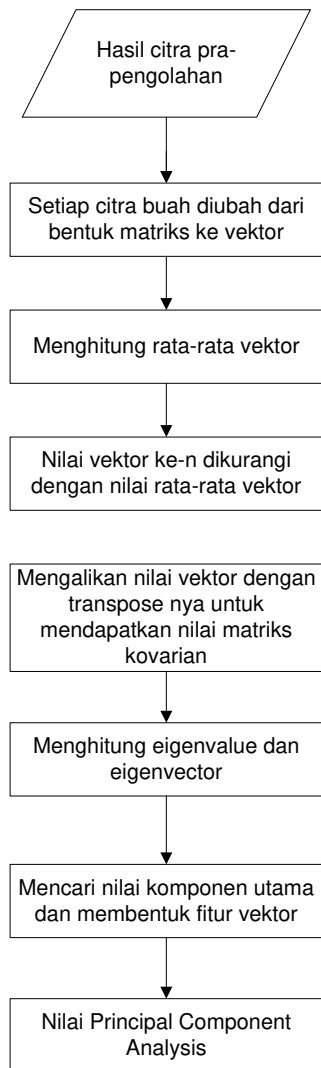
dan *Eigenvector* menggunakan persamaan

$$(\chi \cdot I - C) \vec{x} = 0 \quad \text{dimana } \vec{x} = \text{eigen vector} \quad (5)$$

dengan menggunakan hasil nilai menghitung matrik kovarian, setelah itu dilakukan pengurutan nilai *eigen* dari yang paling besar dan disesuaikan dengan *eigenvector* secara *descending* dari paling yang besar ke yang paling kecil menggunakan persamaan

$$\text{Konstruk} = \chi_i \times \text{MatriksEig} \quad (6)$$

Tahap kelima mencari nilai komponen utama, lalu nilai ekstraksi disimpan kedalam sebuah matrik ekstraksi. Penurunan data set baru ini adalah langkah terakhir dalam proses ekstrak fitur dengan PCA sebelum data tersebut menjadi input dari *minimum distance classifier*. Diagram alir ekstraksi ciri pengenalan citra buah dapat dilihat pada gambar 2.



Gambar 2 Diagram Alir Ekstraksi Ciri Citra Buah

2.3. Minimum Distance Classifier

Setelah nilai ciri/fitur dari masing-masing citra pelatihan telah didapatkan di proses ekstraksi ciri menggunakan metode PCA, tugas *Minimum Distance Classifier* adalah mencari nilai terdekat dari suatu pola citra buah dengan cara membandingkan nilai ciri dari citra yang telah melalui proses pelatihan dengan nilai ciri dari citra pengujian menggunakan persamaan

$$dist(x, z_i) = \sqrt{\sum_{k=1}^m (x - z_i)^2} \quad (7)$$

Hasil paling minimal dari hasil perbandingan atau yang memiliki hasil paling kecil dan mendekati nol dapat dianggap sebagai citra buah yang sama.

III. HASIL PENELITIAN

Sasaran utama digunakannya metode *Principal Component Analysis* (PCA) pada penelitian ini adalah untuk mereduksi dimensi citra sehingga menghasilkan variabel yang lebih sedikit. Sehingga diharapkan waktu yang diperlukan untuk proses pelatihan akan lebih singkat. Hal ini tidak akan berpengaruh terhadap ciri yang dihasilkan oleh PCA karena sifat PCA sendiri adalah mengambil informasi yang paling dominan atau unik dari setiap citra yang disebut dengan eigenvalue. Hal ini telah dibuktikan dengan penggunaan 100 citra buah sebagai dataset yang terbagi atas 50 data latih dan 50 data uji, dengan proses pelatihan 50 citra buah yang memakan waktu selama 27 detik, dan untuk proses pengenalan memakan waktu 0,5 detik per citra buah. Sehingga proses pengenalan dengan menggabungkan *Principal Component Analysis* dengan metode *Minimum Distance Classifier* dapat mempercepat kinerja sistem tanpa mengurangi kehandalan sistem.

Pengujian penelitian ini terbagi menjadi dua, yaitu pengujian dengan citra buah pengenalan yang sama dengan citra buah pelatihan, dan pengujian dengan citra buah pengenalan yang berbeda dengan citra buah pelatihan. Untuk setiap pengujian dengan citra buah pengenalan yang sama dengan citra buah pelatihan, hasil persentase yang didapat adalah sebesar 100%, yang jika dilihat dari sudut pandang PCA, variasi fitur (nilai eigen) yang dihasilkan pada saat proses pelatihan sama dengan variasi fitur (nilai eigen) yang dihasilkan pada saat pengenalan. Namun, persentase pengujian untuk citra buah pengenalan yang berbeda dengan citra buah pelatihan mengalami penurunan sebesar 16%, sehingga didapatkan persentase sebesar 84% karena terdapat 8 citra buah yang gagal dikenali sistem. Faktor-faktor yang mempengaruhi kegagalan tersebut adalah:

1. Nilai fitur (eigenvector) dari citra buah tertentu lebih menyerupai nilai fitur citra buah yang lain.
2. Pengaruh data citra buah input di training set yang lebih menyerupai bentuk buah lain.

Selain itu juga, hasil dari fitur ekstraksi yang tidak signifikan menjadi salah satu penyebab pengenalan menjadi salah. Hal ini terjadi bukan karena kesalahan PCA dalam menentukan ciri, melainkan karena kesalahan peneliti pada saat pengambilan gambar dataset. Kurangnya konsistensi dalam penentuan jarak dari lensa kamera ke objek dan pencahayaan yang berubah-ubah menyebabkan buah tertentu lebih terlihat

Prosiding
ANNUAL RESEARCH SEMINAR 2016
6 Desember 2016, Vol 2 No. 1

ISBN : 979-587-626-0 | UNSRI

<http://ars.ilkom.unsri.ac.id>

seperti buah lain dari sudut pandang PCA, karena dalam penelitian ini PCA berfokus ke piksel di dalam citra buah, bukan ke ciri mendasar buah seperti warna dan teksturnya dan perubahan pencahayaan itu sendiri akan sangat berpengaruh terhadap nilai piksel di dalam citra buah. Terdapat dua solusi untuk menghindari kesalahan pengenalan, yang pertama adalah dengan menambah jumlah data citra di training set, karena sesuai dengan prinsip PCA, semakin banyak variasi citra yang dilatih, kemampuan sistem pengenalan juga akan semakin tinggi. Namun, penambahan jumlah data citra di training set ini dapat memperlambat kerja sistem karena semakin banyak data citra yang dilatih, maka akan semakin banyak pula waktu yang terpakai untuk proses pelatihan, sehingga peneliti tetap mempertahankan jumlah training set yang telah ditentukan sebelumnya. Solusi kedua yaitu dengan menambah jenis fitur yang cocok dengan pengenalan buah, yang dalam hal ini merupakan fitur warna atau tekstur sehingga informasi yang dihasilkan tidak hanya dari piksel di dalam gambar. Solusi kedua ini akan membutuhkan metode lain sebagai pendukung dalam hal mencari fitur warna dan tekstur untuk lebih memudahkan PCA dalam mereduksi fitur-fitur tersebut. Solusi ini diharapkan dapat digunakan untuk penelitian selanjutnya.

IV. KESIMPULAN

Berdasarkan hasil analisis dan implementasi yang telah dilakukan pada penelitian ini, maka dapat diambil kesimpulan sebagai berikut:

1. Hasil dari ekstraksi ciri menggunakan *Principal Component Analysis* (PCA) dapat digunakan dalam penerapan algoritma *Minimum Distance Classifier* untuk pengenalan buah.
2. Jumlah citra latih berbanding lurus dengan kinerja dari sistem pengenalan buah, dalam hal kemampuan melakukan pengenalan terhadap variasi buah yang disajikan ke sistem.
3. Dalam penelitian ini, hasil pengujian dengan menggunakan data latih (*training data set*) mendapatkan keakuratan sebesar 100% sedangkan hasil pengujian dengan menggunakan data uji mendapatkan keakuratan sebesar 84%.
4. *Minimum Distance Classifier* dapat mengenali citra buah dengan baik pada batas perbedaan nilai fitur PCA maksimal 4,39%.
5. Penggabungan metode PCA dan *Minimum Distance Classifier* memberikan solusi yang cepat dan sederhana dengan mempertahankan atau tanpa mengurangi akurasi standar.

DAFTAR PUSTAKA

- [1] Theodoridis, S. and Koutrombas, K. (2008). *Pattern Recognition*. (4th ed.) San Diego: Academic Press.
- [2] Duda, R. O., Hart, P. and Stork, D. (2001). *Pattern Classification*. (2nd ed.) New York: John-Wiley.
- [3] Arivazhagan, S., Shebiah, R. N., Nindhyandhan, S. S. and Ganesan, L. (2010). Fruit Recognition using Color and Texture Features. *Journal of Emerging Trends in Computing and Information Sciences*. 1(2), 90-94.
- [4] Bag, S. and Sanyal, G. (2011). An Efficient Face Recognition Approach using PCA and Minimum Distance Classifier. *International Conference on Image Information Processing*.
- [5] Sarkar, S. and Gupta, S. (2008). A Comprehensive Approach for Fruit Pattern Recognition Under the Framework of Digital Image Processing. *International Journal of Advance Research in Science and Engineering-IJARSE*. 1(2).
- [6] Seng, W. C., and Mirisae, S. H. (2009). A New Method for Fruits Recognition System.