

DOI: doi.org/10.21009/03.1301.FA09

ANALISIS PENGENALAN TINGKAT KEMATANGAN KELAPA MENGGUNAKAN METODE *PRINCIPAL COMPONENT ANALYSIS* (PCA) BERDASARKAN DATA SUARA KETUKAN KELAPA DI RUANG TERBUKA

Yusuf Niko Fitranto^{a)}, Bambang Heru Iswanto^{b)}, Haris Suhendar^{a)}

Program Studi Fisika, FMIPA Universitas Negeri Jakarta, Jl. Rawamangun Muka No. 01, Rawamangun 13220, Indonesia.

Email: ^{a)}unjyusufnikof@gmail.com, ^{b)}bhi@unj.ac.id

Abstrak

Buah kelapa yang dikirim dari petani umumnya memiliki variasi kematangan yang berbeda-beda. Penelitian ini bertujuan untuk mengidentifikasi kematangan buah kelapa berdasarkan analisis fitur akustik dengan menggunakan metode *Principal Component Analysis* (PCA). Metode ini menggabungkan akuisisi data suara ketukan kelapa dengan analisis statistik multivariat untuk mengenali jenis kematangan buah kelapa secara akustik. Eksperimen penelitian melibatkan 40 sampel kelapa dan perekaman suara ketukan buah kelapa dilakukan dengan mengetuk kelapa sebanyak tiga kali dari masing-masing sampel kelapa menggunakan alat pengetuk kelapa di ruang terbuka. Fitur akustik domain waktu dan frekuensi diekstraksi dari sinyal *audio* yang dihasilkan. Selanjutnya, analisis PCA digunakan untuk mengurangi dimensi fitur akustik dan mengidentifikasi pola yang mewakili tingkat kematangan buah kelapa. Hasil visualisasi PCA diperoleh perbedaan fitur akustik dari tingkat kematangan kelapa muda dan tua dapat teridentifikasi. Dari hasil analisis menggunakan PCA diperoleh dua komponen utama pertama yang menjelaskan sekitar 45.32% dan komponen utama kedua menjelaskan sekitar 31 % dari variasi data. Visualisasi data menggunakan *scree plot* menunjukkan kelompok kelapa muda terpisah secara jelas dengan kelompok kelapa tua.

Kata-kata kunci: Identifikasi Suara Ketukan Kelapa, Ruang Terbuka, Kematangan Kelapa, PCA.

Abstract

Coconuts delivered from farmers generally have different variations in maturity. This study aims to identify the maturity of coconut fruit based on acoustic feature analysis using the Principal Component Analysis (PCA) method. This method combines data acquisition of coconut tapping sounds with multivariate statistical analysis to acoustically recognise the type of coconut maturity. The research experiment involved 40 coconut samples and the recording of coconut knocking sound was done by knocking the coconut three times from each coconut sample using a coconut knocking device in an open space. Time and frequency domain acoustic features were extracted from the resulting audio signals. Subsequently, PCA analysis was used to reduce the dimensionality of the acoustic features and identify patterns that represent the ripeness level of the coconut. The results of PCA visualisation obtained differences in acoustic features from young and old coconut maturity levels can be identified. From the results of analysis using PCA, the first two principal components explained about 40.28% and the second principal component explained about 30.07% of the data variation. Visualisation of the data using a *scree plot* shows that the young coconut group is clearly separated from the old coconut group.

Keywords: Coconut Knocking Sound Identification, Open Space, Coconut Maturity, PCA.

PENDAHULUAN

Sektor pertanian dan pangan memiliki peran yang sangat penting dalam perkembangan ekonomi global. Perkembangan dan pertumbuhan populasi manusia membuat kebutuhan sektor pertanian dan pangan semakin meningkat. Sektor pertanian dan pangan tentunya membutuhkan pengolahan dan pengukuran yang cepat dan presisi untuk menghasilkan kualitas yang baik. Seiring berkembangnya teknologi, sektor pertanian dan pangan mendapatkan improvisasi dalam pengukuran dan pengolahan pangan [1]. Bahan pangan yang dibutuhkan salah satunya Kelapa. *Cocos Nucifera* (kelapa) merupakan jenis pangan buah yang populer dan memiliki kandungan gizi tinggi. Kelapa memiliki kepopuleran pada masyarakat asia dan memberikan kontribusi langsung terhadap perkembangan ekonomi dan mata pencaharian daerah tropis [2] [3]. Buah kelapa memiliki struktur yang terdiri dari beberapa bagian: kulit luar (*epicarp*), sabut (*husk*), tempurung (*shell*) dan daging buah (*albumen*) [4]. Tingkat kematangan kelapa dibagi menjadi kelapa muda dan kelapa tua. Kelapa muda memiliki daging buah lunak dan mengandung banyak air sedangkan kelapa tua memiliki daging keras dan tidak ada air. Cara mengklasifikasi atau mengetahui jenis kelapa di kalangan petani atau penjual kelapa dilakukan dengan cara konvensional yakni dengan cara mengetuk kelapa dan mendengarkan suara ketukan yang dihasilkan oleh kelapa [5].

Manusia memiliki kemampuan pendengaran untuk mengidentifikasi suara ketukan kelapa sebagai penanda tingkat kematangan, namun metode ini memiliki risiko tidak akurat serta faktor suara lain diruang terbuka yang dianggap pengganggu (*noise*) menjadi hambatan dalam identifikasi suara ketukan kelapa [6][7]. Suara ketukan kelapa berasal dari getaran yang terdengar oleh telinga manusia [8]. Beberapa pendekatan analisis penelitian suara dilakukan untuk mengenal sinyal dan karakteristik suara [9]. Deskripsi sinyal suara dilakukan dengan memvisualisasikan sinyal *sampling rate* suara dan melibatkan pemahaman fitur-fitur suara, metode matematis dan metode statistik [10]. Metode yang digunakan untuk menganalisis suara yakni domain waktu dan domain frekuensi. Domain waktu dan domain frekuensi menghasilkan nilai numerik fitur-fitur suara yang akan diproses lebih lanjut [11].

METODOLOGI

Penelitian dibagi menjadi beberapa langkah: membuat desain alat pengetuk kelapa, merealisasikan alat pengetuk kelapa, merekam suara pengetukan kelapa di ruang terbuka, mengekstraksi fitur suara kelapa dengan domain waktu dan frekuensi, menerapkan metode PCA dan menganalisis hasil PCA. Data diambil langsung (data primer) dengan 40 sampel kelapa di ruang terbuka dan perekaman suara dilakukan sebanyak 3 kali dengan durasi 1 detik disetiap masing-masing sampel. Dalam klasifikasi jenis kelapa, fitur-fitur domain waktu dan frekuensi dibuat menjadi data numerik serta jenis kelapa muda dan tua dibuat menjadi data kategorik dengan label 0 dan 1. Data numerik ekstraksi fitur dihasilkan 120 data numerik dengan dua label kategori. Ekstraksi fitur domain frekuensi melibatkan fitur *Mel-Frequency Coefficient Cepstral* (MFCC) dan *Power-Neutralized Coefficient Cepstral* (PNCC) serta fitur domain waktu melibatkan *Root Mean Square Energy*, *Zero Crossing Rate* dan *Amplitude Envelope*. Fitur-fitur yang ada akan diseleksi untuk memilih fitur yang informatif dengan metode PCA. Analisis PCA digunakan untuk mengurangi dimensi fitur akustik dan mengidentifikasi pola yang mewakili tingkat kematangan buah kelapa. Hubungan komponen utama dengan fitur suara asli akan menginterpretasikan karakteristik suara yang mewakili tingkat kematangan kelapa.

HASIL DAN PEMBAHASAN

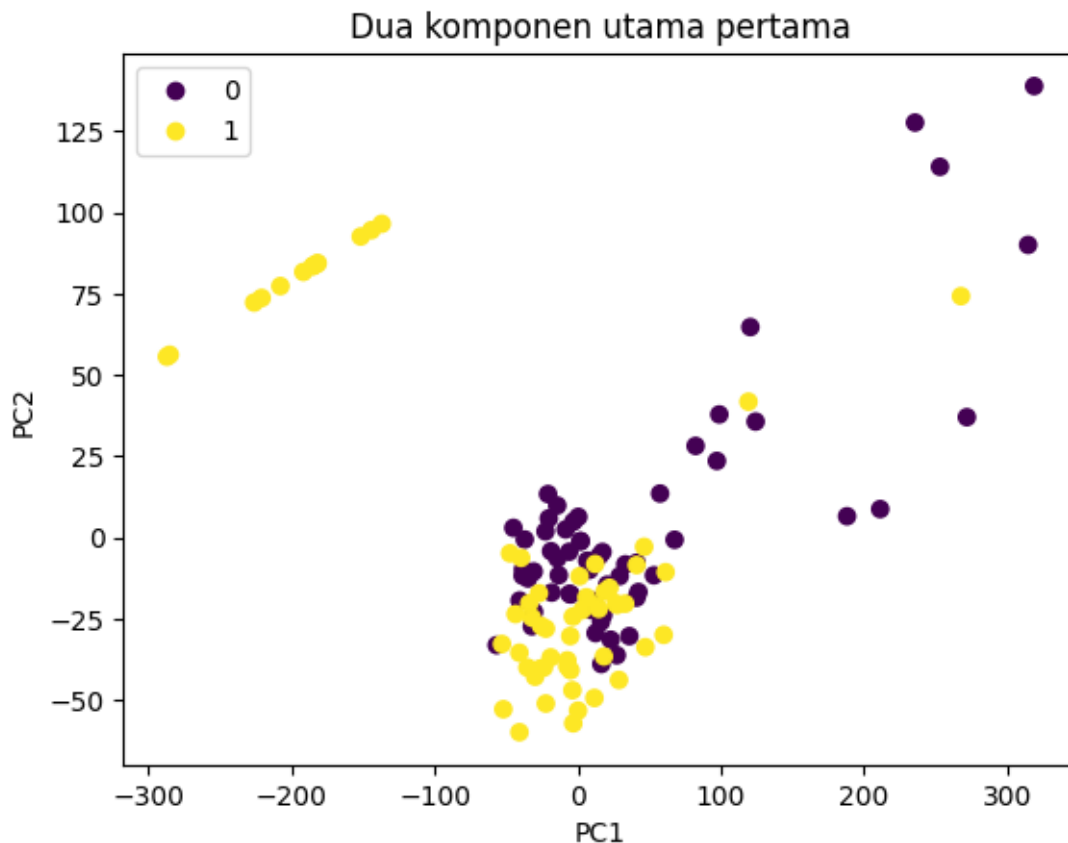
Penelitian dilakukan dengan melakukan ekstraksi fitur data dengan representasi data numerik setiap suara. Ekstraksi fitur menghasilkan 29 fitur yang terdiri dari domain waktu (*zero crossing rate*, *RMS energy*, *amplitude envelope*) dan domain frekuensi (MFCC dan PNCC). Proses ekstraksi fitur menggunakan pemrograman bahasa Python dengan *website* Google Colaboration. Pustaka yang

digunakan yakni *librosa*, *soundfile* dan *matplotlib*. Berikut merupakan tabel hasil ekstraksi fitur dari data *audio* kelapa:

TABEL 1. Hasil ekstraksi fitur domain waktu dan frekuensi pada data *audio* kelapa

Fitur		
MFCC 1	PNCC 1	<i>Amplitude Envelope</i>
MFCC 2	PNCC 2	<i>Zero Crossing Rate</i>
MFCC 3	PNCC 3	<i>RMS Energy</i>
MFCC 4	PNCC 4	
MFCC 5	PNCC 5	
MFCC 6	PNCC 6	
MFCC 7	PNCC 7	
MFCC 8	PNCC 8	
MFCC 9	PNCC 9	
MFCC 10	PNCC 10	
MFCC 11	PNCC 11	
MFCC 12	PNCC 12	
MFCC 13	PNCC 13	

Fitur-fitur diatas dianalisis dengan metode PCA untuk mengurangi dimensi fitur dan mengidentifikasi pola yang mewakili tingkat kematangan buah kelapa dengan menemukan fitur signifikan klasifikasi. Komponen utama pertama (PC1) dan kedua (PC2) menampilkan dua fitur berdasarkan distribusi data. Berikut merupakan visualisasi *score plot* dari kedua fitur:



GAMBAR 1. Distribusi *score plot* data komponen utama fitur

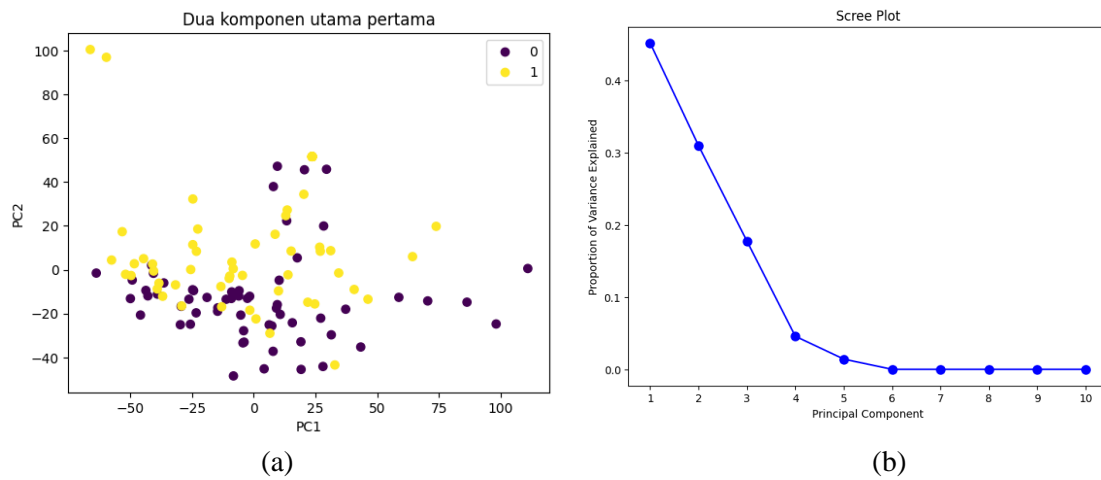
Distribusi *score plot* data diatas terdapat titik yang bertumpuk antara kelapa muda dan kelapa tua. Oleh karena itu, dibutuhkan proses pengolahan statistik data dan seleksi fitur menggunakan uji

anova. Seleksi fitur menunjukkan hasil sebanyak 20 fitur terpilih. Berikut merupakan tabel hasil seleksi fitur terpilih:

TABEL 2. Hasil seleksi fitur domain waktu dan frekuensi pada data *audio* kelapa

Fitur		
MFCC 1	MFCC 6	PNCC 13
MFCC 2	PNCC 2	<i>RMS Energy</i>
MFCC 3	PNCC 3	
MFCC 4	PNCC 8	

Data hasil seleksi fitur diolah dengan PCA untuk melihat visualisasi data dan distribusi data. Berikut merupakan visualisasi *score plot* dan *scree plot* PCA data hasil seleksi fitur:



GAMBAR 2. Distribusi data: (a) *score plot* dan (b) *scree plot* data hasil seleksi fitur

Hasil visualisasi data *score plot* dan *scree plot* menunjukkan seleksi fitur memberikan perubahan distribusi data pada kedua kelas. Distribusi data mulai terpisah secara signifikan berdasarkan kelas masing-masing. Hal ini didukung dengan analisis tabel nilai *eigenvector* sebagai berikut:

TABEL 3. Nilai *eigenvector*

	PC1	PC2	PC3	PC4	PC5
Eigenvalues	1163.29	797.38	457.02	117.69	36.38
Accumulation	1163.29	1960.67	2417.69	2535.38	2571.76
Proportion	45.23%	31.00%	17.77%	4.57%	1.41%

Hasil nilai *eigenvector* menunjukkan bahwa proporsi PC1 memiliki nilai sebesar 45.23% dan PC2 memiliki nilai sebesar 31%. Nilai proporsi PC1 menunjukkan satu komponen representatif dari data asli, sedangkan PC2 dan PC3 memiliki informasi tambahan untuk menjelaskan variasi yang tidak dapat ditangkap oleh PC1. Proporsi PC4 dan PC5 memiliki sedikit informasi tambahan, tetapi tidak terlalu signifikan. Hasil seleksi fitur terpilih dengan *set* fitur MFCC, PNCC dan *RMS Energy* memiliki nilai dominan pada dua komponen utama (PC).

SIMPULAN

Berdasarkan hasil penelitian dapat disimpulkan bahwa penggunaan PCA efektif untuk mengidentifikasi pola tingkat kematangan kelapa dan mengurangi dimensi data menjadi dua komponen utama tanpa kehilangan informasi yang signifikan. Keseluruhan variasi dalam data dapat dijelaskan oleh dua komponen utama dan komponen lainnya sebagai penambahan informasi variasi dalam data. Fitur MFCC, PNCC dan *RMS Energy* dapat mengidentifikasi pola kelas secara jelas dan signifikan. Dari hasil ini penelitian dapat dikembangkan untuk melakukan klasifikasi kematangan buah kelapa menggunakan metode *machine learning* dan *deep learning*.

UCAPAN TERIMAKASIH

Ucapan terima kasih yang sebesar-besarnya saya berikan kepada dosen pembimbing yang telah membimbing dan membantu saya dalam menyusun *paper* ini. Terima kasih yang tulus saya ucapkan kepada orang tua, Mutiah Siregar, sahabat, teman-teman Fisika dan pihak Laboratorium Fisika UNJ yang telah memberikan dukungan dan memfasilitasi penelitian pada *paper* ini. Mohon maaf apabila terdapat kesalahan dan kekurangan dalam penulisan *paper*. Semoga *paper* ini dapat menjadi referensi, memberikan manfaat dan memotivasi pembaca untuk mengembangkan penelitian selanjutnya.

REFERENSI

- [1] K. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, "Machine Learning in Agriculture: a Review," *Sensors*, vol. 18, no. 8, 2018, pp. 2674.
- [2] J. A. Caladcad et al., "Determining Philippine Coconut Maturity Level Using Machine Learning Algorithms Based on Acoustic Signal," *Computers and Electronics in Agriculture*, vol. 172, 2020, pp. 105327.
- [3] Bukhari, D. Leni, Ikbal, Fardinal, and R. Sumiati, "Modifikasi Mesin Pengupas Serabut Kelapa," *Jurnal Surya Teknika*, vol. 9, no. 2, 2022, pp. 450–455.
- [4] G. Mardiatmoko and M. Ariyanti, *Produksi Tanaman Kelapa (Cocos nucifera L.)*, Ambon, Indonesia: Badan Penerbit Fakultas Pertanian Universitas Pattimura, 2018.
- [5] W. A. Nafi'ah and Mitrayana, "Deteksi Frekuensi Akustik pada Buah Kelapa Magelang (Cocos nucifera) Menggunakan Software Spectra PLUS-DT," *Jurnal Fisika Indonesia*, vol. 19, no. 57, 2015, pp. 51–54.
- [6] M. A. Abdillah, "Klasifikasi Tingkat Kematangan Buah Kelapa Menggunakan Deep Learning Berbasis Fitur Akustik," Universitas Negeri Jakarta. 2023.
- [7] R. Y. Sipasulta, A. S. M. Lumenta, and S. R. U. A. Sompie, "Simulasi Sistem Pengacak Sinyal Dengan Metode FFT (Fast Fourier Transform)," *Jurnal Teknik Elektro Dan Komputer*, vol. 3, no. 2, 2014, pp. 1–9.
- [8] D. Halliday, R. Resnick, J. Walker, *Fundamentals of physics*, John Wiley & Sons, 2013.
- [9] T. Theodorou, I. Mporas, N. Fakotakis, "An Overview of Automatic Audio Segmentation," *International Journal of Information Technology and Computer Science (IJITCS)*, vol. 6, no. 11, 2014, pp. 1-9.
- [10] F. Cohen Tenoudji, "Analog and Digital Signal Analysis From Basics to Applications," Springer International Publishing Switzerland, 2016.
- [11] A. G. Jondya, B. H. Iswanto, "Indonesian's Traditional Music Clustering Based on Audio Features," *Procedia Computer Science*, vol. 116, 2017, pp. 174-181.