



ANALISIS PITCH SUARA MENGGUNAKAN PYTHON UNTUK KLASIFIKASI JENIS SUARA VOKAL

Muhammad Hanif Imawan¹, Naufal AthorIQ², Fifi Dewi Ratnasari³

^{1,2}Program Studi Fisika, Fakultas Matematika dan Ilmu Pengetahuan Alam,

Universitas Negeri Semarang, Kampus Sekaran, Gunungpati, Semarang, Jawa Tengah

Email penulis pertama : hanifimawan8@students.unnes.ac.id

Abstract

This study aims to develop a vocal type identification system for male and female voices based on pitch using the probabilistic YIN (pYIN) method implemented in the Python Librosa library. Pitch, or fundamental frequency, serves as a key indicator in classifying vocal types such as tenor, baritone, and bass (male), as well as soprano, mezzo-soprano, and alto (female). The system is designed to read audio files in .wav format, extract pitch in segments, and calculate statistical features such as mean, median, and frequency range to determine vocal classification. Visualizations in the form of time-based curves and pitch histograms provide insights not only into vocal range but also into the dynamics and expressive techniques used by singers. The results indicate that pitch distribution analysis is an effective and informative approach for vocal classification, with potential applications in training, music analysis, and biometric systems.

Keywords: *pitch, pYIN, vocal classification, fundamental frequency, Librosa, voice analysis*

Abstrak

Penelitian ini bertujuan untuk mengembangkan sistem identifikasi jenis suara vokal pria dan wanita berdasarkan pitch menggunakan metode probabilistic YIN (pYIN) dalam pustaka Python Librosa.

Article History

Received: Juni 2025

Reviewed: Juni 2025

Published: Juni 2025

Plagiarism Checker No
235

Prefix DOI :

[10.8734/Kohesi.v1i2.365](https://doi.org/10.8734/Kohesi.v1i2.365)

Copyright : Author

Publish by : Kohesi



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



Pitch atau frekuensi dasar suara merupakan indikator utama dalam mengklasifikasikan jenis vokal seperti tenor, bariton, bass (pria), serta soprano, mezzo-soprano, dan alto (wanita). Sistem dirancang untuk membaca file audio berformat .wav, mengekstrak pitch secara segmental, lalu menghitung statistik seperti rata-rata, median, dan rentang frekuensi guna menentukan klasifikasi jenis suara. Visualisasi dalam bentuk kurva waktu dan histogram pitch memberikan wawasan tidak hanya tentang jangkauan vokal, tetapi juga dinamika dan teknik ekspresif yang digunakan penyanyi. Hasil ini menunjukkan bahwa analisis distribusi pitch merupakan pendekatan efektif dan informatif dalam klasifikasi suara vokal untuk keperluan pelatihan, analisis musik, dan aplikasi biometrik.

Kata kunci: pitch, pYIN, klasifikasi vokal, frekuensi dasar, Librosa, analisis suara

Manusia memiliki karakteristik jenis vokal yang diklasifikasikan berdasarkan rentang nada yang dapat dijangkau secara alami oleh seseorang. Dalam dunia vokal, klasifikasi umum mencakup soprano, mezzo-soprano, dan alto untuk wanita, serta tenor, bariton, dan bass untuk pria. Perbedaan ini terutama ditentukan oleh frekuensi dasar (fundamental frequency/ F_0) dari suara, yang secara langsung dipengaruhi oleh struktur anatomi seperti panjang dan ketebalan pita suara (Sundberg, 1987; Titze, 1994).

Frekuensi dasar merupakan ukuran objektif dari pitch, yaitu persepsi tinggi-rendahnya suara. Meskipun pitch bersifat subjektif, dalam praktik pengolahan sinyal digital, pitch sering digunakan secara sinonim dengan frekuensi dasar karena keterkaitan erat keduanya. Pitch seseorang dapat mencerminkan jenis suara; misalnya, pitch rata-rata pria berkisar 85–180 Hz, sedangkan wanita berkisar 165–255 Hz (Boersma, 1993; Kreiman & Sidtis, 2011).

Proses pengenalan jenis suara secara otomatis umumnya dilakukan dengan mengekstrak pitch dari sinyal audio, menganalisis distribusinya, dan membandingkan hasilnya terhadap



rentang pitch standar berdasarkan jenis kelamin. Parameter seperti nilai rata-rata, median, dan sebaran pitch digunakan untuk menentukan apakah suara termasuk tenor, bariton, soprano, dan sebagainya. Pendekatan ini berguna dalam pelatihan vokal, klasifikasi penyanyi, maupun dalam riset biometrik suara.

Berbagai metode telah dikembangkan untuk ekstraksi pitch, mulai dari pendekatan klasik seperti autocorrelation dan cepstrum, hingga metode modern seperti YIN dan probabilistic YIN (pYIN). YIN dikenal karena akurasi dalam mendeteksi pitch tunggal, sementara pYIN menggabungkan pendekatan probabilistik untuk menangani noise dan variasi suara yang kompleks (De Cheveigné & Kawahara, 2002; Mauch & Dixon, 2014). Pustaka Python Librosa menyediakan implementasi pYIN yang andal dan banyak digunakan dalam penelitian akustik.

Penelitian ini bertujuan mengembangkan sistem berbasis Python yang dapat menganalisis suara dari file .wav, mengekstrak pitch menggunakan pYIN, mengklasifikasikan jenis suara berdasarkan gender dan sebaran pitch, serta memvisualisasikan hasilnya dalam bentuk grafik waktu dan histogram.

METODOLOGI

Penelitian ini menggunakan pendekatan kuantitatif eksperimental dengan metode komputasi berbasis Python. Sistem dibangun dalam bentuk skrip Python yang dapat dijalankan secara interaktif melalui Google Collab atau platform lokal. Berikut adalah tahapan metodologi yang digunakan:

1. Input dan Validasi Data

- Pengguna mengunggah file audio berformat .wav melalui antarmuka unggahan.
 1. Judika - Cinta Karena Cinta
 2. Nadin Amizah - Bertaut
 3. Baby Romeo - Bunga Terakhir
- Sistem memverifikasi bahwa file memiliki ekstensi .wav yang valid.
- Pengguna diminta untuk memasukkan jenis kelamin penyanyi (pria atau wanita).

2. Ekstraksi Pitch



- Durasi total file audio dihitung terlebih dahulu.
- File kemudian dibagi menjadi segmen-segmen berdurasi tetap (misalnya 20 detik).
- Pitch diekstraksi dari setiap segmen menggunakan librosa.pyin dengan rentang frekuensi antara C2 (65 Hz) hingga C6 (1046 Hz).
- Hanya pitch valid (voiced dan bukan NaN) dalam rentang 70–1000 Hz yang dipertahankan.

3. Analisis Statistik Pitch

- Sistem menghitung nilai rata-rata (mean), median, pitch minimum, dan maksimum.
- Pitch disimpan dalam array untuk digunakan dalam visualisasi dan klasifikasi.

4. Klasifikasi Jenis Suara

- Berdasarkan gender yang diinput pengguna, sistem menerapkan logika klasifikasi:
 - Pria : bass, bariton, tenor
 - Wanita: alto, mezzo, soprano
- Rasio distribusi pitch juga diperhitungkan untuk klasifikasi lebih tepat.

5. Visualisasi

- Kurva pitch terhadap waktu ditampilkan untuk 30 detik pertama audio.
- Histogram distribusi pitch digunakan untuk menunjukkan dominasi frekuensi.

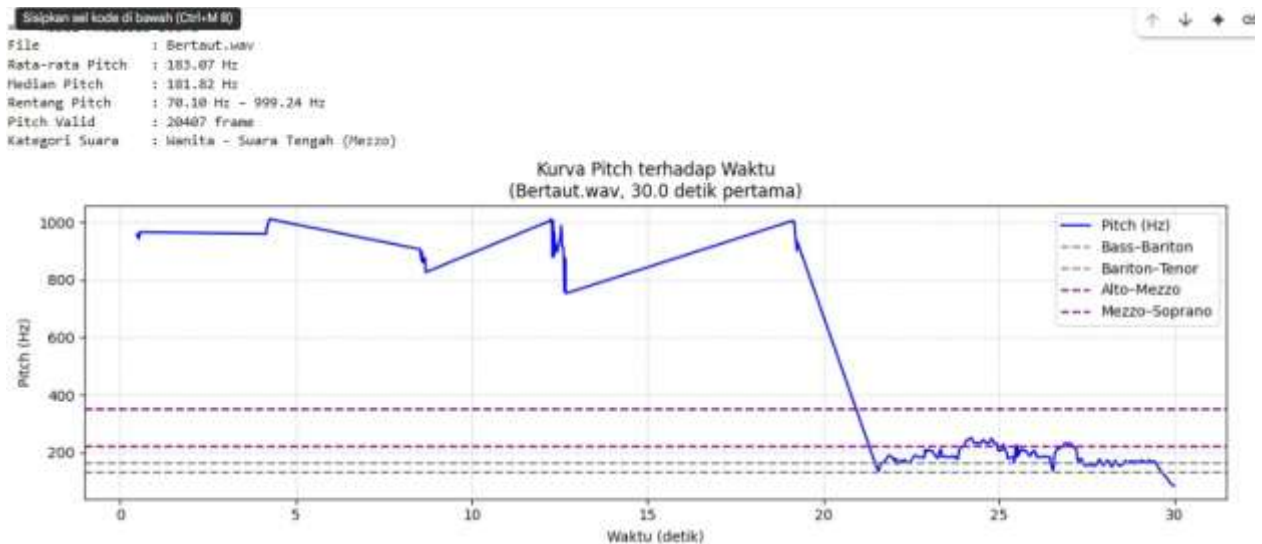
HASIL DAN PEMBAHASAN

Berdasarkan data audio yang telah diinputkan dalam program, proses analisis kemudian dilakukan untuk mengetahui jenis suara vokal yang dinyanyikan dan menampilkan grafik perbandingan waktu dengan pitch 30 detik pertama audio. Berikut adalah hasilnya :



Grafik di atas menunjukkan kurva pitch terhadap waktu untuk 30 detik pertama lagu *Cinta Karena Cinta* yang dinyanyikan oleh Judika. Berdasarkan hasil analisis, rata-rata pitch tercatat sebesar 178,61 Hz, dengan median pitch 155,56 Hz dan rentang pitch mulai dari 78,10 Hz hingga 999,24 Hz. Nilai ini menunjukkan bahwa suara Judika diklasifikasikan sebagai **tenor**, yaitu suara pria dengan jangkauan tinggi. Selama sekitar 25 detik awal, pitch cenderung stabil pada rentang 100–180 Hz, menunjukkan kestabilan dan kontrol vokal yang baik. Namun, mulai detik ke-26 hingga ke-30, terjadi lonjakan drastis melebihi 800 Hz, menandakan bagian klimaks lagu di mana teknik vokal tinggi seperti falsetto kemungkinan digunakan.

Garis bantu pada grafik memperjelas klasifikasi suara dengan batas rentang seperti bass, bariton, alto, mezzo, hingga soprano. Sebagian besar pitch Judika berada di atas garis bariton-tenor (165 Hz), memperkuat identifikasi suara tenor. Puncak pitch yang hampir menyentuh 1000 Hz mencerminkan kemampuan vokal luar biasa yang tidak umum dimiliki oleh penyanyi pria biasa. Visualisasi ini bukan hanya menunjukkan rentang dan dinamika suara, tetapi juga memberikan gambaran konkret tentang fleksibilitas dan kekuatan vokal Judika dalam menyanyikan lagu dengan emosi dan intensitas tinggi.



Grafik menunjukkan distribusi pitch selama 30 detik pertama dari lagu *Bertaut* yang dinyanyikan oleh Nadin Amizah. Rata-rata pitch berada di angka 183,07 Hz dengan median pitch 181,82 Hz, dan rentang pitch yang tercatat adalah antara 70,10 Hz hingga 999,24 Hz. Berdasarkan nilai ini serta distribusi pitch yang ditunjukkan, suara Nadin diklasifikasikan sebagai wanita dengan suara tengah (Mezzo). Pada bagian awal lagu (0–20 detik), kurva pitch menunjukkan dominasi frekuensi tinggi, bahkan beberapa segmen mencapai dan melampaui 900 Hz. Hal ini mengindikasikan bahwa teknik vokal yang digunakan dalam pembukaan lagu memiliki elemen falsetto atau notasi tinggi yang dikontrol dengan teknik khas.

Namun setelah detik ke-20, terjadi penurunan tajam dalam pitch, dan grafik mulai menunjukkan kestabilan pada kisaran 150–250 Hz hingga akhir durasi visualisasi. Hal ini mengisyaratkan pergeseran dinamika vokal dari bagian ekspresif yang tinggi menuju bagian yang lebih lembut dan emosional. Terlihat bahwa suara Nadin Amizah cenderung stabil dan konsisten dalam pengucapan nada. Garis bantu menunjukkan sebagian besar pitch berada di antara batas Alto–Mezzo dan Mezzo–Soprano, tetapi median dan kestabilan frekuensi mendukung klasifikasi Mezzo. Secara keseluruhan, grafik ini mencerminkan teknik vokal Nadin yang tidak hanya ekspresif di nada tinggi, tetapi juga mampu menyeimbangkan dengan kontrol nada tengah secara emosional dan musikal.

*** Hasil Analisis Suara ***
File : Bunga terakhir.wav
Rata-rata Pitch : 139.31 Hz
Median Pitch : 108.11 Hz
Rentang Pitch : 70.10 Hz - 999.24 Hz
Pitch Valid : 12134 frame
Kategori Suara : Pria - Suara Rendah (Bass)



Grafik pitch lagu *Bunga Terakhir* menunjukkan karakteristik suara pria dengan rentang frekuensi dominan pada level rendah. Rata-rata pitch tercatat sebesar 139,31 Hz, dengan median pitch 108,11 Hz, yang mengindikasikan frekuensi dominan berada di bawah wilayah Bariton, bahkan cenderung ke area Bass. Rentang pitch yang terekam adalah antara 70,10 Hz hingga 999,24 Hz, namun sebagian besar nilai pitch berada di bawah garis batas 130 Hz (Bass–Bariton). Oleh karena itu, suara penyanyi diklasifikasikan sebagai Pria – Suara Rendah (Bass). Jumlah frame pitch valid sebanyak 12.134 juga menunjukkan kualitas data yang cukup representatif untuk analisis.

Secara visual, kurva pitch memperlihatkan kestabilan frekuensi rendah terutama dari detik ke-10 hingga detik ke-25. Fluktuasi pitch yang terjadi tetap berada dalam rentang rendah dan menunjukkan teknik vokal yang kuat dalam pengendalian nada rendah. Hal ini memperkuat interpretasi bahwa Bebi Romeo memiliki karakter vokal berat dan dalam yang khas untuk penyanyi pria dengan teknik Bass, menciptakan kesan dalam, melankolis, dan emosional sesuai dengan suasana lagu.



KESIMPULAN

Jenis vokal manusia diklasifikasikan berdasarkan rentang nada alami yang ditentukan oleh frekuensi dasar (pitch) suara dan struktur pita suara. Pria umumnya memiliki pitch 85–180 Hz, wanita 165–255 Hz. Klasifikasi suara otomatis dilakukan dengan mengekstrak dan menganalisis pitch dari rekaman audio, lalu dibandingkan dengan rentang standar. Metode ekstraksi pitch yang akurat seperti pYIN banyak digunakan, termasuk dalam pustaka Python Librosa.

Penelitian ini mengembangkan sistem berbasis Python untuk menganalisis suara .wav, mengekstrak pitch dengan pYIN, mengklasifikasikan jenis suara berdasarkan gender dan pitch, serta memvisualisasikan hasilnya. Ketiga grafik pitch menegaskan bahwa rentang dan distribusi frekuensi vokal dapat digunakan sebagai indikator klasifikasi suara:

- Judika menonjol sebagai tenor dinamis
- Nadin Amizah sebagai mezzo yang ekspresif
- Bebi Romeo sebagai bass dengan kontrol stabil

Visualisasi pitch tidak hanya menggambarkan jangkauan teknis, tetapi juga menyoroti gaya vokal, teknik, serta kekuatan emosional masing-masing penyanyi dalam menyampaikan lagu.

DAFTAR PUSTAKA

1. Boersma, P. (1993). *Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound*. Proceedings of the Institute of Phonetic Sciences, 17, 97–110. PDF
2. De Cheveigné, A., & Kawahara, H. (2002). *YIN, a fundamental frequency estimator for speech and music*. JASA, 111(4), 1917–1930. DOI
3. Kreiman, J., & Sidtis, D. V. (2011). *Foundations of Voice Studies*. Wiley-Blackwell.



4. Mauch, M., & Dixon, S. (2014). *pYIN: A fundamental frequency estimator using probabilistic threshold distributions*. ICASSP, 659–663. DOI
5. Sundberg, J. (1987). *The Science of the Singing Voice*. Northern Illinois University Press.
6. Titze, I. R. (1994). *Principles of Voice Production*. Prentice Hall.