

Identifikasi Alat Musik Tradisional Nusa Tenggara Timur (NTT) Menggunakan Metode Mel Frequency Cepstral Coefficients (MFCC) dan K-Nearest Neighbor (KNN)

Yasinta Anita Kewa Nilan^{a1}, I Ketut Gede Suhartana^{a2}, I Wayan Santiyasa^{a3}, I Gede Surya Rahayuda^{a4}

^aInformatics Departement, Faculty of Mathematics and Natural Science, Udayana University
South Kuta, Badung, Bali, Indonesia

¹nilananita507@email.com

²ikg.suhartana@unud.ac.id

³santiyasa@unud.ac.id

⁴igedesuryarahayuda@unud.ac.id

Abstract

Traditional musical instruments such as gongs and drums from East Nusa Tenggara (NTT) are an important part of the local cultural heritage, but their existence is increasingly threatened due to lack of documentation and difficulties in the sound tuning process. This study presents a digital-based approach to classify the accuracy of the sound of these musical instruments by utilizing the Mel Frequency Cepstral Coefficients (MFCC) feature extraction technique and the K-Nearest Neighbor (KNN) classification algorithm. A total of 238 sound data were recorded and processed to remove noise. The results of the feature extraction were then grouped into three categories: "not right", "almost right", and "already right". The system managed to achieve an accuracy of 92.93% with the best parameter configuration. This solution is expected to help craftsmen in the sound tuning process and contribute to the preservation of regional culture.

Keywords: *Traditional musical instruments, Voice matching, MFCC feature extraction technique and KNN optimization*

1. Pendahuluan

Tenggara Timur (NTT) merupakan salah satu provinsi di Indonesia yang kaya akan warisan budaya, termasuk dalam bidang seni musik tradisional. Salah satu alat musik yang masih digunakan hingga kini dalam berbagai kegiatan adat adalah gong gendang, yang kerap dimainkan dalam upacara pernikahan, penyambutan tamu, dan ritual keagamaan [6]. Alat musik tradisional tidak hanya berfungsi sebagai instrumen hiburan, tetapi juga merepresentasikan nilai-nilai sosial, simbolik, dan spiritual masyarakat lokal [8]. Namun, seiring berjalannya waktu, eksistensi alat musik ini mulai mengalami tantangan. Salah satu permasalahan utama adalah proses pembuatannya yang kompleks serta penyetelan suaranya yang masih sangat bergantung pada keahlian dan intuisi pengrajin [3]. Ketiadaan standar baku dalam penyetelan menyebabkan hasil suara menjadi kurang konsisten, terutama ketika alat musik diproduksi oleh pengrajin berbeda. Di sisi lain, perkembangan teknologi saat ini membuka peluang untuk mengatasi tantangan tersebut melalui pendekatan digital. Pengolahan sinyal digital telah banyak digunakan dalam bidang identifikasi suara, termasuk untuk alat musik tradisional [7]. Salah satu metode yang sering digunakan dalam ekstraksi fitur suara adalah *Mel Frequency Cepstral Coefficients* (MFCC), karena mampu menangkap karakteristik penting dari sinyal audio dan merepresentasikannya secara numerik [1]. Untuk keperluan klasifikasi, algoritma *K-Nearest Neighbor* (KNN) menjadi salah satu metode yang paling umum digunakan karena kesederhanaannya serta kemampuannya dalam mengenali pola data berbasis kemiripan [2]. Selain itu, penelitian sebelumnya juga menunjukkan bahwa MFCC dapat dimodifikasi agar lebih optimal dalam kondisi lingkungan yang memiliki noise [5]. Penggunaan gabungan MFCC dan KNN juga telah diterapkan dalam pengenalan suara alat musik tradisional pada penelitian-penelitian sebelumnya [4]. Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk mengembangkan sistem identifikasi ketepatan bunyi alat musik gong gendang menggunakan metode MFCC dan KNN. Sistem ini diharapkan dapat membantu pengrajin dalam

proses penyetelan suara secara lebih objektif, konsisten, dan efisien, serta mendukung pelestarian budaya lokal melalui integrasi teknologi.

2. Metodologi Penelitian

Metode penelitian ini dirancang untuk menghasilkan sistem yang mampu mengidentifikasi tingkat ketepatan bunyi alat musik tradisional NTT secara efisien dan akurat. Prosesnya mencakup tahapan mulai dari pengumpulan data suara, pengolahan awal untuk pembersihan noise, ekstraksi ciri utama menggunakan MFCC, hingga tahap klasifikasi menggunakan algoritma KNN. Setiap tahapan dilakukan secara sistematis agar hasil identifikasi dapat mendukung kebutuhan pengrajin dalam proses penyetelan.

2.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini diperoleh melalui perekaman langsung alat musik tradisional Gong dan Gendang dari Nusa Tenggara Timur (NTT). Perekaman dilakukan dengan menggunakan handphone 1 dalam lingkungan yang seminimal mungkin dengan gangguan noise eksternal untuk memastikan kualitas suara yang baik. Selain itu, pengaturan perekaman dilakukan dengan memperhatikan aspek teknis seperti jarak mikrofon yang berbeda-beda, dari jarak terdekat hingga jarak jauh serta teknik pukulan alat musiknya juga berbeda setiap kelasnya untuk memperoleh hasil yang optimal. Dalam proses perekaman, setiap alat musik dimainkan oleh pengrajin atau pemain musik yang berpengalaman guna memastikan bunyi yang dihasilkan sesuai dengan standar tradisional. Setiap rekaman dilakukan beberapa kali untuk mendapatkan variasi bunyi yang lebih representatif. Data suara dikategorikan ke dalam tiga label utama yaitu “tidak pas, hampir pas dan sudah pas”. Jumlah data yang didapatkan adalah 238 data audio yang terdiri dari 208 data audio gong 30 data audio gendang.

2.2 Pra-pemrosesan

Pra-pemrosesan mencakup pemangkasan bagian hening, penerapan filter *band-pass* untuk mempertahankan frekuensi utama (80–800 Hz), dan metode pengurangan spektral guna mengurangi noise lingkungan seperti suara angin atau manusia tanpa mengganggu suara utama alat musik.

2.3 Ekstraksi Fitur Menggunakan *Mel Frequency Cepstral Coefficients* (MFCC)

Setelah noise dikurangi, dilakukan ekstraksi fitur menggunakan metode MFCC. Prosesnya ini meliputi tahapan *framing*, *windowing*, *fast fourier transform* (FFT), *Mel Filterbank*, *Log Mel Spectrogram*, *Discrete Cosine Transform* (DCT) dan mendapat vektor MFCC.

a. *Framing*

Sinyal suara yang berupa gelombang kontinu dibagi menjadi potongan-potongan kecil yang disebut *frame*. Panjang *frame* yang digunakan adalah 25 milidetik (ms), dengan *overlap* 50% atau hop size sebesar 10 ms. Pembagian ini dilakukan karena sinyal audio bersifat non-stasioner secara keseluruhan, namun dapat diasumsikan stasioner dalam interval waktu pendek. Jumlah sampel per *frame* dihitung berdasarkan sampling rate, misalnya:

$$\text{Jumlah sampel per frame} = 0,025 \times 16000 = 400 \text{ sampel}$$

b. *Windowing*

Setiap frame kemudian dikalikan dengan fungsi jendela, seperti *Hamming window*, untuk menghaluskan tepi *frame* dan mengurangi efek diskontinuitas. Hal ini mencegah terjadinya kebocoran spektral saat dilakukan transformasi ke domain frekuensi. Rumusnya adalah:

$$w[n] = 0.54 - 0.46 \cdot \cos\left(\frac{2\pi n}{N-1}\right), \quad 0 \leq n \leq N-1 \quad (1)$$

c. *Fast Fourier Transform (FFT)*

Proses ini mengubah sinyal dari domain waktu ke domain frekuensi. FFT menghasilkan spektrum frekuensi dari masing-masing *frame*, yang menunjukkan kandungan frekuensi dan amplitudonya dalam sinyal. Output dari tahap ini berupa *magnitude spectrum*:

$$X[k] = \sum_{n=0}^{N-1} x_w[n] \cdot e^{-j\frac{2\pi}{N}kn}, \quad k = 0, 1, \dots, N-1 \quad (2)$$

Nilai $|X[k]|^2$ kemudian digunakan untuk menghitung spektrum daya dari sinyal.

d. *Mel Filterbank*

Spektrum hasil FFT kemudian dipetakan ke dalam skala *Mel*, yang mendekati persepsi pendengaran manusia terhadap frekuensi. Skala *Mel* lebih sensitif terhadap perubahan pada frekuensi rendah. Digunakan sejumlah filter segitiga yang saling tumpang tindih untuk menangkap energi frekuensi dalam rentang-rentang tertentu.

$$\text{Mel}(f) = 2595 \cdot \log_{10} \left(1 + \frac{f}{700} \right) \quad (3)$$

e. *Log Mel Spectrogram*

Untuk mendekati cara pendengaran manusia terhadap intensitas suara, dilakukan transformasi logaritmik terhadap energi filter *Mel*:

$$\hat{E}_m = \log(E_m + \varepsilon) \quad (4)$$

dengan ε adalah bilangan kecil untuk menghindari nilai tak terhingga saat energi sangat kecil atau nol.

f. *Discrete Cosine Transform (DCT)*

Langkah akhir adalah menerapkan *Discrete Cosine Transform (DCT)* pada log-energi filter *Mel* untuk menghasilkan koefisien MFCC. DCT bertujuan untuk mengurangi korelasi antar fitur dan mengonsolidasikan informasi ke dalam beberapa koefisien dominan pertama:

$$c_n = \sum_{m=1}^M \hat{E}_m \cdot \cos \left[\frac{\pi n}{M} \left(m - \frac{1}{2} \right) \right], \quad n = 1, 2, \dots, L \quad (5)$$

- M: jumlah *filter Mel*,
- L : jumlah koefisien MFCC yang diambil (biasanya 12–20 pertama),
- Cn: koefisien MFCC ke-n.

g. *Agregasi Koefisien*

Karena setiap file audio terdiri dari banyak *frame* dan masing-masing *frame* menghasilkan vektor MFCC, dilakukan perhitungan rata-rata untuk setiap koefisien agar diperoleh satu vektor fitur tunggal per file. Vektor ini kemudian digunakan sebagai input untuk proses klasifikasi dengan KNN.

2.4 Normalisasi dan Kategori Data

Setelah proses ekstraksi fitur selesai, nilai-nilai koefisien MFCC dinormalisasi ke dalam rentang 0 hingga 1 menggunakan metode *MinMaxScaler*. Normalisasi ini bertujuan agar setiap fitur memiliki skala yang setara sehingga tidak mendominasi perhitungan jarak pada algoritma KNN. Selanjutnya, dilakukan proses kategorisasi terhadap label data. Kategori label ditentukan berdasarkan tingkat ketepatan bunyi hasil penilaian pengrajin, dengan tiga kelas: 0 untuk “tidak pas”, 1 untuk “hampir pas”, dan 2 untuk “sudah pas”. Proses ini memastikan bahwa data siap digunakan dalam tahap klasifikasi.

2.5 Pembagian Data

Sebelum dilakukan klasifikasi, data dibagi menjadi dua bagian, yaitu 80% sebagai data latih dan 20% sebagai data uji. Pembagian ini dilakukan secara *stratified* untuk menjaga proporsi distribusi kelas tetap seimbang pada kedua bagian.

2.6 Klasifikasi Menggunakan K-Nearest Neighbor (KNN)

Proses klasifikasi dilakukan menggunakan algoritma K-Nearest Neighbor (KNN). Algoritma ini mengklasifikasikan data uji berdasarkan kedekatan jaraknya terhadap data latih menggunakan metrik Euclidean. Nilai K ditentukan melalui pengujian beberapa parameter, dan hasil terbaik diperoleh saat $K = 7$. Hasil klasifikasi menunjukkan performa model yang tinggi dalam membedakan kategori ketepatan bunyi, yaitu “tidak pas”, “hampir pas”, dan “sudah pas”.

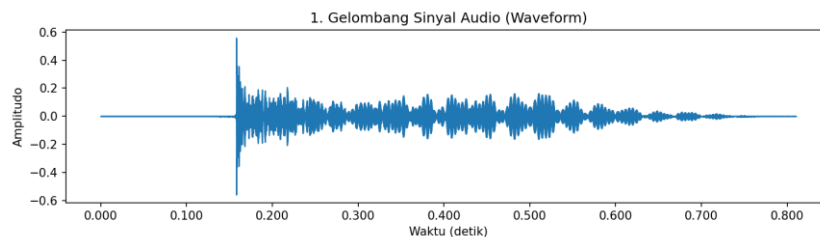
3. Hasil dan Pembahasan

Hasil penelitian menunjukkan bahwa proses ekstraksi fitur menggunakan *Mel Frequency Cepstral Coefficients* (MFCC) berhasil menangkap karakteristik penting dari sinyal suara alat musik tradisional Gong dan Gendang. Proses ini menghasilkan 60 koefisien MFCC per *frame* yang kemudian dirata-ratakan untuk mewakili satu file audio.). Berikut adalah hasil dari setiap tahapan pada salah satu data alat musik gendang:

3.1. Hasil Ekstraksi Fitur Menggunakan *Mel Frequency Cepstral Coefficients* (MFCC)

a. Representasi Gelombang Sinyal Audio

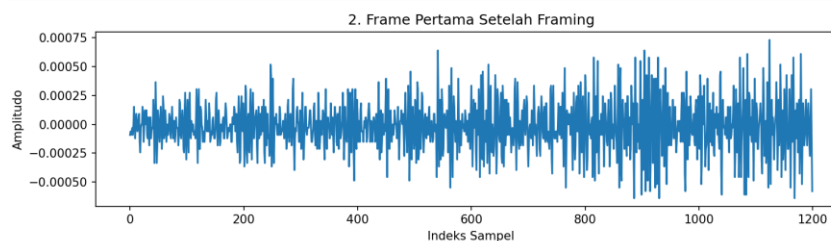
Gambar 1 menunjukkan representasi sinyal audio dalam domain waktu. Lonjakan amplitudo pada grafik menunjukkan adanya suara keras atau impuls dominan dari instrumen. Amplitudo bervariasi antara -0.5 hingga 0.5, yang mencerminkan kekuatan bunyi alat musik yang terekam secara langsung.



Gambar 1. *Waveform* Gelombang Sinyal Audio

b. Framing

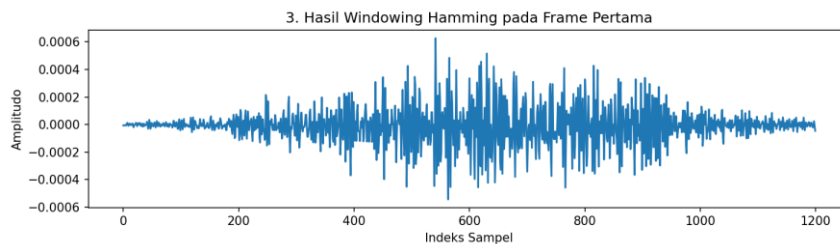
Gambar 2 memperlihatkan salah satu frame pertama dari sinyal audio setelah proses framing. Sinyal dibagi menjadi *frame* pendek berdurasi 25 ms dengan pergeseran 10 ms agar dapat dianalisis secara lokal. Ini penting karena karakteristik suara berubah seiring waktu.



Gambar 2. *Frame* Pertama Proses Framing

c. *Windowing*

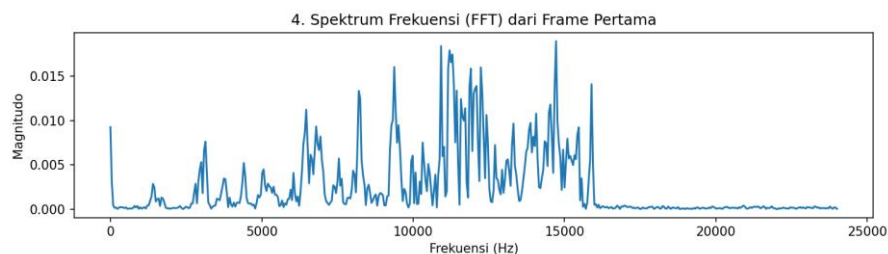
Gambar 3 menunjukkan hasil aplikasi fungsi *Hamming window* pada *frame* pertama. Tampak bahwa sinyal amplitudonya perlahan mengecil ke nol di awal dan akhir *frame* (indeks sampel sekitar 0-200 dan 1000-1200), sementara amplitudo di bagian tengah (sekitar indeks 400-800) tetap tinggi. Proses ini meminimalkan efek transisi tajam di batas *frame*.



Gambar 3. *Frame* Pertama Proses *Windowing*

d. *Fast Fourier Transform* (FFT)

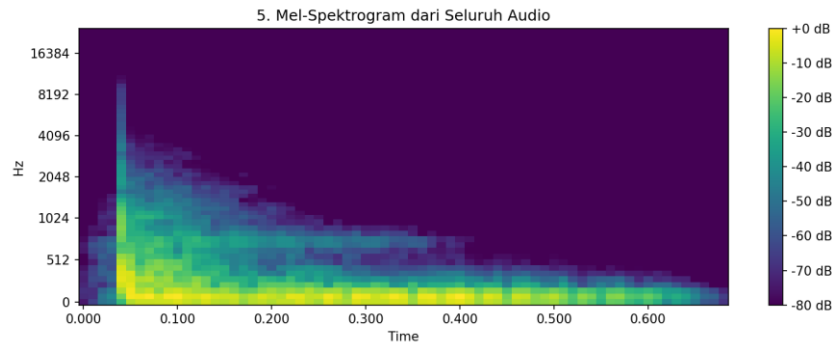
Gambar 4 menunjukkan spektrum hasil Fast Fourier Transform (FFT) dari *frame* pertama. Rentang frekuensi dari 0 Hz hingga sekitar 24.000 Hz, mencakup sebagian besar spektrum audio yang dapat didengar manusia. Adapun puncak-puncak frekuensi yang menonjol dan terpisah-pisah, terutama dari frekuensi rendah hingga sekitar 16.000 Hz. Puncak-puncak ini, seperti yang terlihat di sekitar 1.000 Hz, 6.000 Hz, 10.000 Hz, dan 15.000 Hz, mengindikasikan keberadaan frekuensi-frekuensi dominan yang merupakan bagian integral dari sinyal asli, kemungkinan besar adalah harmonik (kelipatan frekuensi dasar) yang membentuk karakter suara, seperti timbre atau warna suara pada alat musik tersebut.



Gambar 4. *Frame* Pertama Proses FFT

e. *Mel Filterbank*

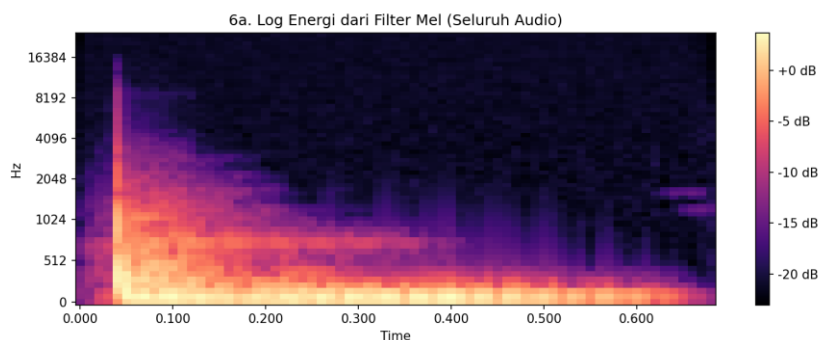
Gambar 5 menunjukkan hasil dari proses Mel Filterbank. Kekuatan suara (ditunjukkan dengan warna) tersebar di berbagai frekuensi dan berubah seiring waktu. Pada sumbu vertikal menggunakan skala Mel (seperti 0 Hz hingga 16384 Hz), yang disesuaikan agar mirip dengan cara telinga manusia mendengar. Frekuensi rendah "diregangkan" lebih lebar, sedangkan frekuensi tinggi "dipadatkan". Pada awal suara (sekitar 0.000 hingga 0.100 detik), terlihat energi yang sangat kuat (kuning terang) di frekuensi rendah di bawah 512 Hz yang kemudian cepat menyebar ke frekuensi lebih tinggi, menunjukkan "*attack*" atau permulaan suara yang tiba-tiba. Setelah itu (0.100 hingga 0.700 detik), energi suara berangsur melemah (warna memudar), terutama di frekuensi tinggi yang cepat menghilang, sementara energi terkuat bertahan lebih lama di frekuensi rendah hingga menengah (di bawah 2048 Hz), menunjukkan karakteristik "*decay*" atau gema yang perlahan menghilang.



Gambar 5. Mel-Spektrogram Seluruh Audio

f. *Log Mel Spectrogram*

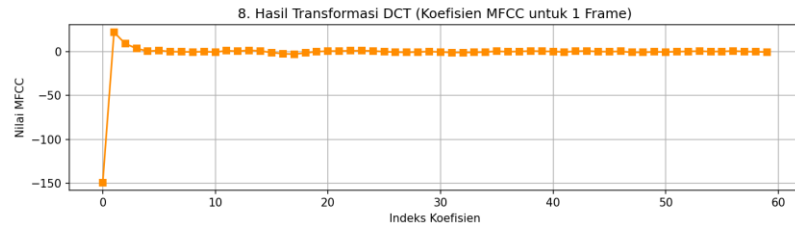
Gambar 6 log energi hasil dari *Mel Filterbank* menunjukkan bahwa adanya bunyi yang sangat kuat dan kaya nada di awal rekaman (terlihat dari area kuning terang di bagian kiri). Dan jika dilihat lebih detile ada pola yang muncul di sekitar 0 - 512 Hz (Frekuensi Rendah), 512 Hz – 2048 Hz (Frekuensi Menengah), dan diatas 2048 Hz (Frekuensi Tinggi) . Pada rentang frekuensi renda terlihat ada jejak energi yang bertahan cukup lama, hingga sekitar 0.600 detik, meskipun dengan intensitas yang semakin melemah (berwarna oranye/merah). Ini mungkin menunjukkan adanya komponen "bass" atau nada dasar yang bergaung lebih lama. Lalu pada rentang frekuensi tinggi menunjukkan energi yang cukup kuat di awal dan perlahan mereda. Ada beberapa "garis" horizontal samar berwarna oranye yang menunjukkan adanya frekuensi-frekuensi tertentu yang mungkin terus bergetar atau beresonansi. Dan pada rentang frekuensi tinggi tampaknya mereda lebih cepat dibandingkan frekuensi rendah. Area kuning/oranye di bagian atas spectrogram ini jauh lebih singkat. Ini wajar, karena nada-nada tinggi cenderung lebih cepat menghilang atau diredam dalam banyak suara alami.



Gambar 6. Filter Mel Seluruh Audio

g. *Discrete Cosine Transform (DCT)*

Gambar 7 menunjukkan bahwa terdapat 60 koefisien pada sumbu horisontal dan sekitar -150 hingga 20 nilai MFCC. Kurva oranye pada grafik ini menunjukkan nilai-nilai koefisien MFCC untuk satu potongan suara. Koefisien pertama (indeks 0, sekitar -150) mencerminkan energi suara secara keseluruhan, sementara koefisien awal lainnya (indeks 1 hingga sekitar 10-15) sangat penting karena mereka menggambarkan "warna" atau karakter unik (timbre) dari suara tersebut. Koefisien yang lebih tinggi (di atas indeks 15) nilainya cenderung nol dan sering diabaikan karena kurang memberikan informasi penting tentang suara.



Gambar 7. Discrete Cosine Transform (DCT)

h. Hasil MFCC

Gambar 8 menunjukkan tabel hasil akhir ekstraksi fitur MFCC. Setiap kolom mewakili satu nilai koefisien MFCC dari satu frame suara. Data ini kemudian dirata-ratakan dan digunakan sebagai input dalam proses klasifikasi KNN.

File_Nam/Class	MFCC 1	MFCC 2	MFCC 3	MFCC 4	MFCC 5	MFCC 6	MFCC 7	MFCC 8	MFCC 9	MFCC 10	MFCC 11	MFCC 12	MFCC 13	MFCC 14	MFCC 15	MFCC 16	MFCC 17
Gazaki-1 tidak pas	-158.734	5.809001	7.392215	0.088945	0.75393	2.440822	2.116757	1.612263	1.920217	0.904111	-0.6361	-1.76349	-1.30277	-0.78639	-0.87484	-0.78875	-0.35857
Gazaki-1 tidak pas	-153.38	9.938182	8.174727	-1.11084	-0.00219	1.390171	0.898158	0.102995	0.538243	0.063804	-0.77966	-1.29661	-0.58731	-0.28607	-0.80328	-1.37179	-1.19628
Gazaki-1 tidak pas	-149.106	12.79631	7.71591	-1.40425	-0.43276	0.269249	-0.26051	-0.78133	-0.3052	-0.45725	-0.48227	-0.48129	0.061187	-0.01123	-0.82899	-1.10675	-0.79031
Gazaki-1 tidak pas	-145.187	14.96323	6.649841	-1.62788	-0.4049	-0.16248	-0.61102	-0.82704	-0.36397	-0.88276	-0.41367	-0.09721	0.10071	-0.04479	-0.21748	-0.13546	-0.23751
Gazaki-1 tidak pas	-141.2	16.78025	5.40222	-0.91583	0.38334	-0.23776	-0.59743	-0.6241	-0.13125	-0.78301	0.225229	0.694379	0.087437	-0.19263	0.151251	0.158447	0.381317
Gazaki-1 tidak pas	-124.303	19.52615	-0.25401	0.293385	-0.34656	0.464511	-0.64364	-1.35192	0.091681	-1.21444	-1.28495	0.98326	-0.1032	-0.77352	0.555679	-0.30303	0.27975
Gazaki-1 tidak pas	-105.782	23.31818	-5.60022	2.225018	-1.32863	2.060327	-1.42551	-1.02286	0.069885	-0.99758	-1.99043	0.972202	0.024441	-1.14099	0.767899	-1.38414	0.31622
Gazaki-1 tidak pas	-91.4721	26.37817	-10.8247	3.387154	-2.1601	3.451202	-1.65366	-1.10453	0.082918	-1.38925	-1.7018	1.112498	0.273714	-0.89601	0.905431	-1.62931	0.454702
Gazaki-1 tidak pas	-80.5031	28.37315	-16.1476	3.392301	-2.58197	4.496669	-2.08138	-2.45028	-0.29658	-2.11814	-1.8266	0.925813	0.685871	-0.41425	0.978164	-1.43494	0.624229
Gazaki-1 tidak pas	-82.6846	28.80461	-18.1004	1.90248	-1.73715	3.82205	-1.5545	-3.62027	-0.30595	-2.54015	-0.75599	1.182252	1.0266	0.442862	0.797539	-0.48509	0.878909
Gazaki-1 tidak pas	-85.603	27.97732	-18.0145	-0.51774	-0.76313	1.896063	-1.48932	-5.32228	-0.55777	-2.58324	-0.63004	1.588908	1.235815	1.025847	0.434947	0.231458	1.217849

Gambar 8. Hasil MFCC

3.2. Hasil Klasifikasi Menggunakan K-Nearest Neighbors (KNN)

Setelah proses ekstraksi ciri, langkah berikutnya adalah klasifikasi menggunakan algoritma *K-Nearest Neighbors* (KNN).

a. Panjang *Frame* dan Pergeseran *Frame* Terhadap Akurasi

Tabel 1. Pengujian Panjang *Frame* dan Pergeseran *Frame* Terhadap Akurasi

Panjang <i>Frame</i>	Panjang Pergeseran <i>Frame</i>	Jumlah Koefisien MFCC	Jumlah K	Akurasi Data <i>Training</i>	Akurasi Data <i>Testing</i>
0,015 detik	0,05 detik	8	7	91.50%	84.99%
	0,01 detik	8	7	94.92%	90.69%
0,025 detik	0,05 detik	8	7	94.86%	89.67%
	0,01 detik	8	7	92.93%	90.28%

Berdasarkan tabel 1 hasil pengujian, saya memilih kombinasi panjang *frame* 0,025 detik dan pergeseran *frame* 0,01 detik sebagai parameter terbaik untuk model. Pemilihan ini didasarkan pada kestabilan model, yang terlihat dari selisih akurasi antara data *training* dan *testing* yang paling kecil, yaitu 2,65% (*training* 92,93% dan *testing* 90,28%). Selisih kecil ini menunjukkan kemampuan model untuk mengeneralisasi dengan baik tanpa *overfitting*. Meskipun kombinasi lain memiliki akurasi *testing* sedikit lebih tinggi, seperti 0,015 detik dengan akurasi 90,69%, selisih akurasi yang lebih besar pada kombinasi tersebut menunjukkan potensi *overfitting*. Oleh karena itu, kombinasi 0,025 detik dan 0,01 detik dianggap lebih optimal untuk performa yang stabil dan dapat diandalkan pada data baru.

b. Jumlah Koefisien MFCC Terhadap Akurasi

Tabel 2. Pengujian Jumlah Koefisien MFCC Terhadap Akurasi

Panjang <i>Frame</i>	Panjang Pergeseran <i>Frame</i>	Jumlah Koefisien MFCC	Jumlah K	Akurasi Data <i>Training</i>	Akurasi Data <i>Testing</i>
0,025 detik	0,01 detik	8	7	92.93%	90.28%

Panjang Frame	Panjang Pergeseran Frame	Jumlah Koefisien MFCC	Jumlah K	Akurasi Data Training	Akurasi Data Testing
0,025 detik	0,01 detik	10	7	91,90%	84,12%
0,025 detik	0,01 detik	13	7	96,45%	92.54%
0,025 detik	0,01 detik	30	7	98.51%	91,23%
0,025 detik	0,01 detik	40	7	98.51%	93.97%
0,025 detik	0,01 detik	60	7	99,88%	92,26%

Berdasarkan tabel 2 hasil pengujian, jumlah koefisien MFCC yang optimal adalah 8, karena memberikan kestabilan dan kemampuan generalisasi yang baik. Konfigurasi ini menghasilkan akurasi training 92,93% dan testing 90,28%, dengan selisih hanya 2,65%. Selisih kecil ini menunjukkan bahwa model tidak mengalami *overfitting* dan dapat mengenali pola dengan konsisten. Meskipun ada jumlah koefisien lain yang menunjukkan akurasi testing lebih tinggi, seperti 40 (93,97%) dan 13 (92,54%), selisih akurasi pada konfigurasi tersebut lebih besar dari 3%, yang menandakan risiko *overfitting*. Oleh karena itu, MFCC 8 dipilih sebagai parameter terbaik karena model ini stabil dan efektif untuk data baru.

c. Nilai K Terhadap Akurasi

Tabel 3. Pengujian Nilai K Terhadap Akurasi

Panjang Frame	Panjang Pergeseran Frame	Jumlah Koefisien MFCC	Jumlah K	Akurasi Data Training	Akurasi Data Testing
0,025 detik	0,01 detik	8	7	92.93%	90.28%
0,025 detik	0,01 detik	10	7	91,90%	84,12%
0,025 detik	0,01 detik	13	7	96,45%	92.54%
0,025 detik	0,01 detik	30	7	98.51%	91,23%
0,025 detik	0,01 detik	40	7	98.51%	93.97%
0,025 detik	0,01 detik	60	7	99,88%	92,26%

Berdasarkan tabel 3 hasil pengujian untuk menentukan nilai K terbaik pada algoritma *K-Nearest Neighbor* (KNN), evaluasi dilakukan terhadap berbagai nilai K untuk mengukur dampaknya pada performa model. Hasil menunjukkan bahwa K = 7 adalah parameter terbaik, dengan akurasi *training* 92,93% dan *testing* 90,28%, serta selisih akurasi hanya 2,65%. Meskipun bukan akurasi *tertinggi*, selisih kecil ini menunjukkan stabilitas model dan kemampuannya untuk menggeneralisasi dengan baik pada data baru. Oleh karena itu, K = 7 dipilih sebagai pilihan optimal, karena menawarkan keseimbangan antara kinerja tinggi dan risiko *overfitting* yang rendah.

4. Kesimpulan

Berdasarkan penelitian mengenai identifikasi alat musik tradisional Nusa Tenggara Timur (NTT) menggunakan metode *Mel Frequency Cepstral Coefficients* (MFCC) dan *K-Nearest Neighbor* (KNN), dapat disimpulkan bahwa pemilihan parameter yang tepat pada tahap ekstraksi ciri dan klasifikasi sangat memengaruhi akurasi dan kestabilan model. Penggunaan panjang *frame* 0,025 detik dan pergeseran *frame* 0,01 detik terbukti optimal dengan akurasi *training* 92,93% dan *testing* 90,28%, menunjukkan model dapat menggeneralisasi data dengan baik tanpa *overfitting*. Selain itu, jumlah koefisien MFCC juga berperan penting, di mana penggunaan 8 koefisien memberikan keseimbangan antara akurasi dan kestabilan model, sementara jumlah koefisien yang lebih besar cenderung menyebabkan *overfitting*. Pada algoritma KNN, nilai K=7 menghasilkan performa terbaik dan paling stabil, ditunjukkan dengan perbedaan akurasi training dan testing yang kecil. Secara keseluruhan, kombinasi parameter tersebut mendukung keberhasilan sistem dalam mengenali kualitas nada alat musik tradisional NTT dan berpotensi membantu para pengrajin dalam proses penyetelan.

Daftar Pustaka

- [1] A. Solihin, D. I. Mulyana, dan M. B. Yell, "Klasifikasi Jenis Alat Musik Tradisional Papua menggunakan Metode Transfer Learning dan Data Augmentasi," *Semantics Scholar*, 2022. [Online]. Tersedia: <https://www.semanticscholar.org/paper/Klasifikasi-Jenis-Alat-Musik-Tradisional-Papua-dan-Solihin-Mulyana/8e44cce9e5e1bd3c6e836a58edf1de86b6ee40c3> [Diakses: 2 Mei 2023].
- [2] A. Anggoro, S. Herdjunto, dan R. Hidayat, "MFCC dan KNN untuk Pengenalan Suara Artikulasi P," *Jurnal Avitec*, 2020. [Online]. Tersedia: <https://ejournals.itda.ac.id/index.php/avitec/article/view/605> [Diakses: 1 Mei 2023].
- [3] A. F. Ryamizard, B. Hidayat, dan S. Saidah, "Deteksi Nada Tunggal Alat Musik Kecapi Bugis Makassar Menggunakan Metode Mel Frequency Cepstral Coefficient (MFCC) dan Klasifikasi K-Nearest Neighbour (KNN)," *Jurnal Engineering*, 2018. [Online]. Tersedia: <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/7702> [Diakses: 1 Mei 2023].
- [4] S. Hussein, "Mengenal K-Nearest Neighbor: Algoritma Populer untuk Machine Learning," *Geospasialis.com*, 2021. [Online]. Tersedia: <https://geospasialis.com/k-nearest-neighbor/> [Diakses: 2 Mei 2023].
- [5] Y. R. Prayogi, "Modifikasi Metode MFCC untuk Identifikasi Pembicara di Lingkungan Ber-Noise," *Jointecs*, 2019. [Online]. Tersedia: <https://publishing-widyagama.ac.id/ejournal-v2/index.php/jointecs/article/view/999> [Diakses: 2 Mei 2023].
- [6] J. Anna, "5 Jenis Alat Musik Berdasarkan Sumber Bunyinya," *Adjar Grid*, 2022. [Online]. Tersedia: <https://adjar.grid.id/read/543487604/5-jenis-alat-musik-berdasarkan-sumber-bunyinya> [Diakses: 3 Mei 2023].
- [7] B. Update, "Pengertian Identifikasi Beserta Contoh dan Prosesnya," *Umparan.com*, 2021. [Online]. Tersedia: <https://umparan.com/berita-update/pengertian-identifikasi-beserta-contoh-dan-prosesnya-1wqwdBE3gcz> [Diakses: 3 Mei 2023].
- [8] Kuncoro, "Pengertian dan Jenis Alat Musik Tradisional," *Matsansaga.com*, 2020. [Online]. Tersedia: <https://www.matsansaga.com/2020/01/pengertian-dan-jenis-alat-musik-tradisional.html> [Diakses: 5 Mei 2023].

This page is intentionally left blank.