

# Classification of Bird Sounds Using the Mel-Frequency Cepstral Coefficient (MFCC) and K-Nearest Neighbor (KNN) Methods

Anak Agung Gde Ramananda Kartikeya Pattraksha<sup>a1</sup>, I Wayan Supriana<sup>a2</sup>, I Made Widiartha<sup>a3</sup>, Luh Arida Ayu Rahning Putri<sup>a4</sup>, Ida Bagus Gede Dwidasmara<sup>a5</sup>, I Dewa Made Bayu Atmaja Darmawan<sup>a6</sup>

<sup>a</sup>Informatics Departement, Udayana University  
Jimbaran, Bali, Indonesia

<sup>1</sup>kartikeyaramananda@email.com

<sup>2</sup>wayan.supriana@unud.ac.id

<sup>3</sup>madewidiartha@unud.ac.id

<sup>4</sup>rahningputri@unud.ac.id

<sup>5</sup>dwidasmara@unud.ac.id

<sup>6</sup>dewabayu@unud.ac.id

## Abstract

*With the rapid development of this technology, of course, it can provide benefits to the wider community, one of which is to provide convenience in meeting the needs of an agency or group, one of which is forest rangers and animal supervisors who can assist in classifying the types of birds or poultry, because birds have voices. which differ by type. For example, in the West Bali National Park itself, there are several types of bird species that are there, for example, the Java Dederuk bird or birds belonging to the Columbidae species. In this case technology plays an important role as a medium for information exchange regarding information about birds and can easily determine the type of bird. So, to fulfill this, we need a system that can classify bird species based on their sound. In this study, the system built uses the MFCC method for feature extraction. This method was chosen because the success rate in speech recognition using MFCC feature extraction is higher than feature extraction using FFT. And for the classification stage using the KNN method. This method was chosen because KNN is easy to represent compared to other methods. From the results of the Classification of Bird Sounds Using the Mel-Frequency Cepstral Coefficient (MFCC) and K-Nearest Neighbor (KNN) methods, it can be concluded that the resulting K-Nearest Neighbor model is able to classify bird species based on their sound properly because after a single data classification with data outside of the training data, the system is able to classify it very accurately, the accuracy percentage is 80%. And in the test results using K-fold Cross Validation, the best K value is obtained in the K-Nearest Neighbor model, namely K = 1 with 77% accuracy.*

**Keywords:** Mel Frequency Cepstral Coefficient, K-Nearest Neighbor, Birds

## 1. Pendahuluan

Dengan pesatnya perkembangan teknologi ini, tentunya dapat memberikan manfaat bagi masyarakat luas, salah satunya mempermudah pemenuhan kebutuhan suatu instansi atau kelompok, salah satunya jagawana dan pengawas satwa yang dapat membantu dalam mengklasifikasikan jenis burung atau unggas, karena burung memiliki suara. yang berbeda menurut jenisnya. Misalnya saja di Taman Nasional Bali Barat sendiri terdapat beberapa jenis jenis burung yang ada disana, misalnya burung Dederuk Jawa atau burung yang termasuk dalam jenis Columbidae. Dalam hal ini, teknologi berperan penting sebagai sarana pertukaran informasi terkait informasi burung dan dapat dengan mudah menentukan jenis burung tersebut. Untuk mencapai hal tersebut diperlukan suatu sistem yang dapat mengklasifikasikan jenis burung berdasarkan suaranya.

Oleh karena itu pada penelitian ini sistem yang dibangun menggunakan metode MFCC untuk ekstraksi fitur. Metode ini dipilih karena tingkat keberhasilan pengenalan suara menggunakan ekstraksi fitur MFCC lebih tinggi dibandingkan dengan ekstraksi fitur menggunakan FFT [1]. Dan untuk tahap klasifikasi menggunakan metode KNN. Metode ini dipilih karena KNN mudah direpresentasikan dibandingkan dengan metode lainnya [2].

**2. Metode Penelitian**

**2.1. Data dan Metode Pengumpulan Data**

Data untuk penelitian ini diambil dari burung-burung di Bali. Jumlah data yang digunakan dalam penelitian ini adalah 50 data suara burung, dimana terdapat 5 kelas dan untuk setiap kelas terdapat 10 data suara. Famili burung yang dipilih untuk klasifikasi ini adalah perkutut jawa, gagak hutan, cerek besar, dan dara laut batu. Data suara burung diperoleh dari database burung dunia avibase.bsc-eoc.org. Format data yang digunakan adalah data audio dalam format .wav. Jenis burung yang akan digunakan dalam penelitian ini dapat dilihat pada Tabel 1.

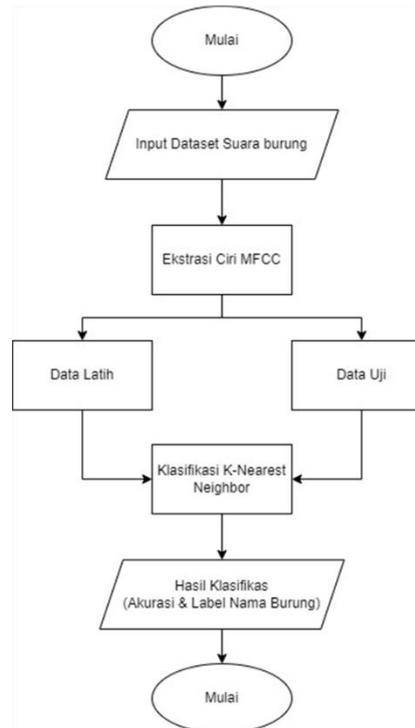
**Tabel 1. Bird Data**

<b>Nama ilmiah</b>	<b>Nama Burung</b>	<b>Famili</b>	<b>Label (.wav)</b>
Geopelia striata	perkutut jawa	columbidae	748014
Geopelia striata	perkutut jawa	columbidae	738793
Geopelia striata	perkutut jawa	columbidae	618483
Geopelia striata	perkutut jawa	columbidae	616403
Geopelia striata	perkutut jawa	columbidae	598171
Geopelia striata	perkutut jawa	columbidae	588080
Geopelia striata	perkutut jawa	columbidae	576417
Geopelia striata	perkutut jawa	columbidae	764718
Geopelia striata	perkutut jawa	columbidae	756428
Geopelia striata	perkutut jawa	columbidae	713711
Corvus enca	gagak hutan	corvidae	657295
Corvus enca	gagak hutan	corvidae	628125
Corvus enca	gagak hutan	corvidae	616145
Corvus enca	gagak hutan	corvidae	616142
Corvus enca	gagak hutan	corvidae	616141
Corvus enca	gagak hutan	corvidae	614886
Corvus enca	gagak hutan	corvidae	614885
Corvus enca	gagak hutan	corvidae	614882
Corvus enca	gagak hutan	corvidae	614881
Corvus enca	gagak hutan	corvidae	614880
Pluvialis squatarola	cerek besar	charadriidae	775072
Pluvialis squatarola	cerek besar	charadriidae	775071
Pluvialis squatarola	cerek besar	charadriidae	775069
Pluvialis squatarola	cerek besar	charadriidae	775068
Pluvialis squatarola	cerek besar	charadriidae	774421
Pluvialis squatarola	cerek besar	charadriidae	774420
Pluvialis squatarola	cerek besar	charadriidae	774419
Pluvialis squatarola	cerek besar	charadriidae	774418
Pluvialis squatarola	cerek besar	charadriidae	771426

Pluvialis squatarola	cerek besar	charadriidae	770143
Onychoprion anaethetus	dara laut batu	Laridae	612658
Onychoprion anaethetus	dara laut batu	Laridae	589066
Onychoprion anaethetus	dara laut batu	Laridae	740587
Onychoprion anaethetus	dara laut batu	Laridae	735967
Onychoprion anaethetus	dara laut batu	Laridae	589062
Onychoprion anaethetus	dara laut batu	Laridae	735240
Onychoprion anaethetus	dara laut batu	Laridae	612657
Onychoprion anaethetus	dara laut batu	Laridae	589067
Onychoprion anaethetus	dara laut batu	Laridae	539164
Onychoprion anaethetus	dara laut batu	Laridae	539136
Haliastur indus	Elang Bondol	Accipitridae	743675
Haliastur indus	Elang Bondol	Accipitridae	736313
Haliastur indus	Elang Bondol	Accipitridae	702713
Haliastur indus	Elang Bondol	Accipitridae	685003
Haliastur indus	Elang Bondol	Accipitridae	663577
Haliastur indus	Elang Bondol	Accipitridae	663576
Haliastur indus	Elang Bondol	Accipitridae	663575
Haliastur indus	Elang Bondol	Accipitridae	663573
Haliastur indus	Elang Bondol	Accipitridae	644542
Haliastur indus	Elang Bondol	Accipitridae	578793

## 2.2. Desain Penelitian

Jalannya penelitian tentang klasifikasi suara dengan ekstraksi fitur MFCC menggunakan KNN diamati dengan flowchart. Flowchart digunakan untuk memudahkan penggambaran aliran informasi dan aliran penelitian. Alur desain penelitian ini dapat dilihat pada Gambar 1.



**Gambar 1.** Flowchart Research Design

Pada tahap klasifikasi dengan K-Nearest Neighbor, data dengan fitur terpilih akan dilatih terlebih dahulu. Kemudian data yang telah dilatih akan digunakan untuk memprediksi data uji masukan. Kemudian hasil prediksi ditampilkan dengan menggunakan metode K-Nearest Neighbor.

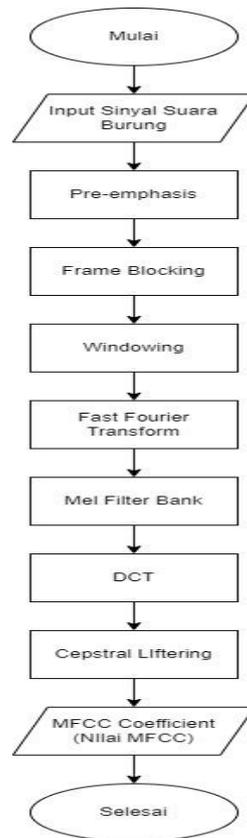
Setelah melakukan prediksi terhadap data uji, selanjutnya dilakukan pengujian terhadap program. Untuk pengujiannya sendiri digunakan metode K-Fold Cross Validation. Setelah itu akan muncul berapa akurasi yang dihasilkan oleh program tersebut. Kemudian hasil perhitungan KNN akan mendeteksi apakah suara burung termasuk kelas perkutut jawa, gagak hutan, cerek besar, dan dara laut batu.

### 2.3. Pre-Processing

Sebelum dilakukan ekstraksi fitur menggunakan Mel Frequency Cepstral Coefficient, file suara burung dalam format .wav akan diproses terlebih dahulu yaitu pembacaan data menggunakan tools SciPy dimana file suara berupa sinyal analog akan diubah menjadi bentuk digital.

### 2.4. Feature Ekstraksi Using Mel Frequency Cepstral Coefficient

Setelah melalui proses preprocessing, langkah selanjutnya adalah ekstraksi fitur menggunakan MFCC. Ekstraksi fitur MFCC merupakan adaptasi terhadap sistem pendengaran manusia, di mana sinyal audio difilter secara linear untuk frekuensi di bawah 1000 Hz dan secara logis untuk frekuensi di bawah 1000 Hz [3]. Tahapan MFCC dapat dijelaskan pada Gambar 2.



**Gambar 2.** Flowchart MFCC

Penjelasan tahapan ekstraksi ciri dengan MFCC pada penelitian ini adalah sebagai berikut:

**a. Pre-Emphasis**

Pre-Emphasis adalah langkah pertama dalam ekstraksi fitur MFCC. Proses ini akan mempertahankan frekuensi tinggi dalam spektrum yang biasanya hilang atau dihilangkan dalam proses produksi suara. Pre-phasis bertujuan untuk mengurangi rasio noise terhadap sinyal sehingga dapat meningkatkan kualitas suatu sinyal dan menyeimbangkan spektrum suara yang disuarakan [3].

**b. Frame Blocking**

Frame Blocking adalah tahap dimana sinyal suara disegmentasi menjadi beberapa frame. Secara umum, dalam proses pemblokiran frame, setiap frame berdurasi 20-25 milidetik dengan porsi tumpang tindih ( $M$ ) antara satu frame dengan frame lainnya sebesar 30-50% dari panjang frame. Panjang frame ( $N$ ) yang digunakan dalam proses ini sangat berpengaruh. untuk sukses dalam analisis spektral [3]. Semakin panjang ukuran frame yang digunakan, semakin banyak resolusi frekuensi yang akan dituju, namun hal ini akan mempengaruhi resolusi waktu yang dihasilkan.

**c. Windowing**

Pada proses Windowing dilakukan tahap pembobotan data untuk setiap frame yang dibentuk pada proses sebelumnya dengan menggunakan fungsi window [4]. Windowing ini dilakukan untuk mengurangi gap setelah proses frame blocking.

**d. Fast Fourier Transform**

Fast Fourier Transform (FFT) merupakan pengembangan dari Discrete Fourier Transform (DFT) yang bertujuan untuk mengubah sinyal dari domain waktu ke domain frekuensi [5]. Pada langkah FFT ini, masing-masing  $N$  sampel dikonversi dari domain waktu ke domain frekuensi.

Setelah mendapatkan nilai FFT, nilai ini digunakan untuk menghitung densitas spektral energi. Densitas spektral energi ini akan digunakan untuk memetakan nilai FFT ke bank filter yang akan digunakan pada tahap selanjutnya.

**e. Filter Bank**

Selanjutnya dilakukan perhitungan bank filter terhadap sinyal data yang diperoleh dari proses FFT sebelumnya. Untuk menyederhanakan nilai ini diubah menjadi satuan dB, sehingga akan dihasilkan nilai bank filter signal.

**f. Discrete Cosinus Transform**

Pada tahap ini, nilai spektrum mel pada domain frekuensi akan diubah menjadi domain waktu untuk mendapatkan nilai koefisien.

**g. Cepstral Liftering**

Pada proses Cepstral lifting ini merupakan teknik yang digunakan untuk memperkecil sensitivitas koefisien cepstral yang dihasilkan dari tahapan utama dalam ekstraksi fitur menggunakan MFCC. Proses ini berfungsi untuk meningkatkan kualitas pengenalan suara.

Setelah melakukan perhitungan yang sama untuk setiap nilai koefisien cepstral, maka akan didapatkan nilai koefisien cepstral terfilter. Nilai ini nantinya akan digunakan sebagai dataset pada tahap klasifikasi.

**2.5. Pembagian Data**

Proses selanjutnya adalah memisahkan dataset menjadi data latih dan data uji. Data pelatihan adalah data yang diperiksa model, sedangkan data uji adalah data yang digunakan untuk mengetahui seberapa baik kinerja model pada data yang tidak terlihat. Scikit-learn memiliki fungsi yang disebut "train\_test\_split" yang memudahkan pemisahan dataset menjadi data pelatihan dan data pengujian.

**2.6. Klasifikasi Menggunakan K-Nearest Neighbor Method**

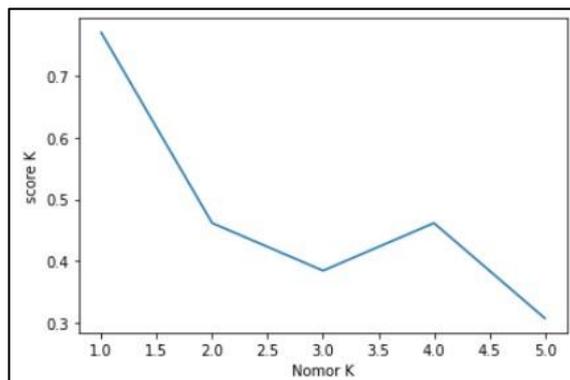
Metode K-nearest neighbor menggunakan semua data uji pada data, setelah itu sistem menentukan data latih dan uji. Setelah Anda menentukan data pelatihan dan pengujian, Anda dapat menemukan jarak minimum dan maksimum dalam perhitungan data pelatihan. Kemudian cari jarak minimum dari data uji ke data pelatihan. Setelah diperoleh jarak antara data latih dan data uji, ditentukan kelas sesuai dengan jarak minimum interval yang telah ditentukan. Setelah itu, hasil klasifikasi ditampilkan dalam bentuk akurasi. Alur metode KNN dapat dilihat pada Gambar 3.



**Gambar 3.** Flowchart KNN

### 3. Hasil dan Pembahasan

Pada Gambar 4 merupakan grafik yang diperoleh dari nilai K yang telah diuji dengan validasi K-Fold cross. Dari hasil pengujian dapat disimpulkan bahwa nilai K berpengaruh terhadap akurasi sistem. Dari hasil pengujian diperoleh nilai K terbaik yaitu K = 1 dengan akurasi 77%.



**Gambar 4.** The results of the K-Nearest Neighbor test with K-Fold Cross Validation

Setelah mendapatkan nilai K terbaik yaitu K = 1, dilakukan laporan Klasifikasi dan didapatkan hasil precision, recall dan f1-score. Dapat dilihat pada Gambar 5.

	precision	recall	f1-score	support
cerek besar	1.00	1.00	1.00	2
dara laut batu	0.67	0.67	0.67	3
elang bondol	0.67	1.00	0.80	2
gagak hutan	0.75	1.00	0.86	3
perkutut jawa	1.00	0.33	0.50	3
accuracy			0.77	13
macro avg	0.82	0.80	0.76	13
weighted avg	0.81	0.77	0.74	13

**Gambar 5.** Classification Report

#### 3.1. Pengujian Sistem dengan Inputan Audio Baru Burung

Pada tahap ini dilakukan pengujian terhadap masukan data audio burung yang baru. Data baru yang akan diuji adalah 5 kelas dengan masing-masing kelas berisi 5 data audio burung baru. Setelah dilakukan pengujian terhadap 25 data audio tunggal untuk burung baru, diperoleh hasil klasifikasi seperti pada Tabel 2.

**Tabel 2.** Classification Results New bird audio input

Nama ilmiah	Famili	Nama File (.wav)	Nama Burung	Hasil Klasifikasi (Nama Burung)
Geopelia striata	columbidae	197147	perkutut jawa	gagak hutan
Geopelia striata	columbidae	665873	perkutut jawa	dara laut batu
Geopelia striata	columbidae	769104	perkutut jawa	perkutut jawa
Geopelia striata	columbidae	767816	perkutut jawa	perkutut jawa
Geopelia striata	columbidae	576417	perkutut jawa	gagak hutan
Corvus enca	corvidae	614879	gagak hutan	gagak hutan
Corvus enca	corvidae	614878	gagak hutan	gagak hutan
Corvus enca	corvidae	105944	gagak hutan	gagak hutan
Corvus enca	corvidae	769106	gagak hutan	gagak hutan
Corvus enca	corvidae	614858	gagak hutan	gagak hutan
Pluvialis squatarola	charadriidae	593996	cerek besar	cerek besar

Pluvialis squatarola	charadriidae	770143	cerek besar	cerek besar
Pluvialis squatarola	charadriidae	774418	cerek besar	cerek besar
Pluvialis squatarola	charadriidae	774420	cerek besar	cerek besar
Pluvialis squatarola	charadriidae	775068	cerek besar	cerek besar
Onychoprion anaethetus	Laridae	769107	dara laut batu	dara laut batu
Onychoprion anaethetus	Laridae	767680	dara laut batu	dara laut batu
Onychoprion anaethetus	Laridae	759450	dara laut batu	elang bondol
Onychoprion anaethetus	Laridae	37739	dara laut batu	dara laut batu
Onychoprion anaethetus	Laridae	251964	dara laut batu	elang bondol
Haliastur indus	Accipitridae	743675	elang bondol	elang bondol
Haliastur indus	Accipitridae	736313	elang bondol	elang bondol
Haliastur indus	Accipitridae	702713	elang bondol	elang bondol
Haliastur indus	Accipitridae	685003	elang bondol	elang bondol
Haliastur indus	Accipitridae	663577	elang bondol	elang bondol

Berdasarkan Tabel 2. Apakah dapat diperoleh hasil pada pengujian data tunggal dimana pada data perkutut Jawa terdapat 2 data yang berhasil diklasifikasikan dengan benar dari 5 data perkutut Jawa. Pada data gagak hutan terdapat 5 data yang berhasil diklasifikasikan dengan benar dari 5 data gagak hutan. pada data burung besar terdapat 5 data yang sudah terklasifikasi dengan benar dari 5 data burung besar. pada data merpati karang terdapat 3 data yang berhasil diklasifikasikan dengan benar dari 5 data merpati karang. pada data elang botak terdapat 5 data yang berhasil diklasifikasikan dengan benar dari 5 data elang botak. aDengan demikian diperoleh hasil dari total 20 data burung yang diuji, 20 data burung berhasil diklasifikasikan dengan benar. Yang mana jika dihitung akurasi sebagai berikut:

$$accuracy = \frac{20}{25} \times 100\% = 80\%$$

Jadi akurasi yang didapat pada uji input data single audio baru diluar data training yang didapatkan adalah akurasi sebesar 80%.

#### 4. Kesimpulan

Dari hasil penelitian yang dilakukan dengan menggunakan metode Mel-Frequency Cepstral Coefficient (MFCC) dan K-Nearest Neighbor (KNN) klasifikasi suara burung, dapat disimpulkan sebagai berikut:

- Model K-nearest neighbor yang dihasilkan cukup baik dalam mengklasifikasikan jenis burung berdasarkan suara karena setelah mengklasifikasikan data individual dengan data di luar data training, sistem dapat mengklasifikasikannya dengan akurasi sebesar 80%
- Pada hasil pengujian yang diperoleh dengan K-fold cross-validation didapatkan nilai K terbaik pada model K-nearest neighbor yaitu H. K = 1 dengan akurasi 77%.

### References

- [1] F. Budiman, M. A. Nursyeha, M. Rivai, and Suwito, "Pengenalan Suara Burung Menggunakan Mel Frequency Cepstrum Coefficient dan Jaringan Syaraf Tiruan Pada Sistem Pengusir Hama Burung" *Jurnal Nasional Teknik Elektro*, vol. 5. No.1, 64-72, 2016
- [2] M. I. Sikki, "Pengenalan wajah menggunakan k-nearest neighbour dengan praproses transformasi wavelet" *paradigma*, Vol.10. No. 2, 159-172, 2009
- [3] M. A. Hilmi, "Metoda Mel Frequency Cepstrum Coefficients (MFCC) untuk Mengenali Ucapan pada Bahasa Indonesia" *Jurnal sains dan teknologi informasi*, Vol. 1. No.1. 22-31, 2012
- [4] T. Chamidy, "Metoda Mel Frequency Cepstrum Coefficients (MFCC) pada Klasifikasi Hidden Markov Model (HMM) untuk Kata Arabic pada penutur Indonesia" *Jurnal MATICS*, Vol. 8. No.1. 36-39, 2016
- [5] H. Heriyanto, S. Hartati, and A. E. Putra, "Ekstraksi Ciri Mel Frequency Cepstral Coefficient (Mfcc) Dan Rerata Coefficient Untuk Pengecekan Bacaan Al-Qur'an" *Telematika: Jurnal Informatika dan Teknologi Informasi*, Vol. 15 No. 2, 99-108, 2018

*This page is intentionally left blank.*