



KLASIFIKASI SPESIES BURUNG BERDASARKAN
SUARA MENGGUNAKAN METODE
CONVOLUTIONAL NEURAL NETWORK DAN
SUPPORT VECTOR MECHINE

SEMINAR BIDANG KAJIAN

AHMAD HIDAYAT
99219022

PROGRAM DOKTOR TEKNOLOGI
INFORMASI UNIVERSITAS GUNADARMA
SEPTEMBER 2022

Daftar Isi

Daftar Isi	i
1 Pendahuluan	1
1.1 Latar Belakang	1
1.2 Batasan dan Tujuan	4
1.3 Kontribusi	4
2 Tinjauan Pustaka	5
2.1 Suara Burung	5
2.2 Deep Learning.	5
2.3 MFCC (Mel Frequency Cepstral Coefficients)	6
2.4 Spektrogram	7
2.5 Arsitektur CNN	8
2.6 Support Vector Machine (SVM)	9
2.7 Perbandingan Tinjauan	10
3 Metodologi	12
3.1 Gambaran Umum Penelitian.	12
3.2 Rencana Kerja	13
Daftar Pustaka	14

Bab 1

Pendahuluan

1.1 Latar Belakang

Antara tahun 1973 (tahun dibuatnya undang-undang tentang jenis hewan langka) sampai tahun 1990, tercatat 26 jenis hewan punah (Upadhyay, Rai and Iyengar 2001). Salah satu jenis tersebut merupakan jenis burung yang pada tahun 2018 tercatat bahwa 8 burung punah selama satu dekade terakhir (Gibbens, 2018). Beragam faktor dapat menyebabkan jenis burung punah, salah satunya adalah hilangnya habitat burung. Mendeteksi keberadaan burung melalui habitatnya merupakan salah satu langkah awal untuk menentukan upaya konservasi yang perlu dilakukan. Oleh sebab itu, diperlukan metode untuk mengklasifikasikan jenis burung melalui habitatnya sebagai salah satu upaya konservasi jenis burung yang perlahan mulai mengalami kepunahan.

Burung memiliki habitat yang beranekaragam, antara lain hutan, perkebunan, padang rumput, perairan dan lainlain. Penyebaran jenis burung dipengaruhi oleh kesesuaian lingkungan tempat hidup burung yang meliputi adaptasi burung terhadap perubahan lingkungan, kompetisi, dan seleksi alam (Alikodra 2002). Beberapa habitat burung dapat dijangkau oleh manusia, namun kemampuan manusia dalam menjangkau habitat burung masih terbatas. Suara burung merupakan salah satu penanda yang khas untuk menandakan adanya kehidupan burung. Dengan teknologi informasi dan komunikasi, manusia dapat mengenali jenis burung yang ditemuinya dengan suara khas tersebut. Saat ini penelitian mengenai klasifikasi jenis burung berdasarkan suara semakin meningkat, antara lain klasifikasi burung berdasarkan suaranya (Raghuram, Chavan, Belur, & Koolagudi, 2016), klasifikasi suara burung menggunakan CNN (Tóth & Czeba, 2016) dan Identifikasi Jenis Burung Berdasarkan Suara (Elias Sprengel, Jaggi, Kilcher, & Hofmann, 2016).

Dalam dekade terakhir, klasifikasi suara burung semakin mendapat perhatian karena penurunan populasi di seluruh dunia. Oleh karena itu, semakin penting untuk melindungi keanekaragaman hayati burung, di mana pemantauan populasi burung adalah langkah pertama untuk perlindungan. Metode tradisional untuk memantau

burung memakan waktu dan biaya. Kemajuan terbaru dalam jaringan sensor akustik nirkabel dan teknik pembelajaran mendalam memberikan cara baru untuk memantau populasi hewan mengandalkan jaringan sensor nirkabel, suara burung dapat dikumpulkan secara terus menerus di lingkungan terbuka, yang kemudian dapat digunakan untuk memantau populasi burung. Namun, berbagai sumber suara dan rendah rasio signal-to-noise dari rekaman yang dikumpulkan menjadi masalah penting, terutama ketika membangun sistem klasifikasi suara burung otomatis yang kuat. (J. Xie, K. Hu, dkk, 2019).

Generasi suara dan sistem persepsi hewan telah berevolusi untuk membantu mereka bertahan hidup di lingkungan mereka. Dari perspektif evolusi, suara yang disengaja yang dihasilkan oleh hewan harus berbeda dari suara acak lingkungan. Beberapa hewan memiliki keistimewaan kemampuan sensorik, seperti penglihatan, pemandangan, perasaan, dan kesadaran akan perubahan alam yang dibandingkan kepada manusia. Suara binatang dapat bermanfaat bagi manusia dalam hal keamanan, prediksi bencana alam, dan interaksi intim jika kita mampu mengenalnya dengan baik. (Emanuella Cecilia Tania, dkk, 2021).

Saat ini, penggunaan berbagai sensor untuk memantau populasi spesies hewan di lingkungan alam menjadi semakin penting. Salah satu tantangan utama yang harus dihadapi para peneliti, adalah untuk mendeteksi segmen peristiwa suara yang berbeda dalam rekaman besar yang diperoleh dari sensor yang beroperasi terus-menerus yang ditempatkan di lapangan. Karena ini adalah tugas yang sangat memakan waktu, akan lebih mudah untuk mengembangkan teknologi baru untuk mengotomatisasi proses deteksi dan klasifikasi.

Selama beberapa tahun terakhir, beberapa penelitian telah dilakukan di bidang deteksi dan klasifikasi suara otomatis di lingkungan luar ruangan. Beberapa penulis telah memfokuskan penelitian mereka dalam studi tentang suara lingkungan (baik alam dan manusia yang dihasilkan), sementara yang lain telah mencoba tugas yang lebih spesifik dari deteksi dan klasifikasi berbagai spesies hewan. Banyak algoritma pembelajaran mesin yang diawasi telah digunakan dalam pengembangan sistem deteksi dan klasifikasi. Membandingkan algoritma tradisional seperti Decision Trees, Support Vector Machine (SVM), k-Nearest Neighbors dan Hidden Markov Models (HMM), secara umum, SVM mencapai hasil terbaik.

Sejak pengenalan Neural Networks (NN) untuk pengenalan pola, NN telah

mengungguli hasil yang diperoleh dengan algoritma tradisional. Misalnya, dalam sistem untuk klasifikasi suara perkotaan kinerja SVM jika dibandingkan dengan konfigurasi yang berbeda dari jaringan saraf seperti jaringan saraf berulang (RNN), jaringan saraf dalam (DNN) dan Jaringan Saraf Konvolusi (CNN), maka menggunakan CNN atau DNN lebih baik dari pada menggunakan SVM atau RNN. Adavanne mengusulkan arsitektur yang menggunakan jaringan saraf berulang dua arah untuk deteksi burung dan memperoleh hasil yang mirip dengan CNN. (Diez Gaspon Itxasne, dkk, 2019).

Support Vector Machine (SVM) merupakan salah satu metode yang telah banyak digunakan untuk pengenalan pola. Pada proses audio, SVM digunakan untuk pembagian segmen fonetis, pengenalan suara dan klasifikasi suara secara umum. Salah satu keunggulan dari SVM adalah akurasi yang tinggi dan lebih unggul dalam menggeneralisasikan dibanding metode lain. Metode machine learning ini bertujuan untuk menemukan hyperplane terbaik yang memisahkan dua buah class pada ruang input. (Miftahul Afida Annisa, 2020).

Klasifikasi pola berdasarkan jaringan dalam mengungguli metode konvensional dalam banyak tugas. Namun, jika database untuk pelatihan menunjukkan representasi internal yang tidak memiliki kejelasan substansial untuk kelas yang berbeda, jaringan dianggap bahwa pembelajaran pada dasarnya gagal. Kegagalan tersebut terbukti ketika akurasi turun tajam dalam eksperimen yang melakukan tugas klasifikasi di mana suara binatang diamati serupa. Untuk mengatasi dan memperbaiki masalah pembelajaran, makalah ini mengusulkan pendekatan baru yang terdiri dari kombinasi beberapa CNN yang masing-masing dilatih secara terpisah untuk menghasilkan fitur tingkat menengah menurut setiap kelas dan kemudian digabungkan menjadi unit CNN gabungan dengan SVM untuk klasifikasi keseluruhan. Untuk percobaan, terlebih dahulu dibuat database suara hewan yang mencakup 3 kelas dengan 102 spesies. Dari hasil eksperimen menggunakan database, metode yang diusulkan terbukti lebih unggul dari metode konvensional yang menonjol. (Kyungdeuk Ko, dkk, 2018).

Berdasarkan penelitian terdahulu, tujuan dari penelitian ini adalah mengembangkan dan mengimplementasikan model atau metode klasifikasi suara pada obyek suara burung untuk menentukan spesies dari model Convolutional Neural Network (CNN) dan Support Vector Mechine (SVM).

1.2 Batasan dan Tujuan

Batasan masalah pada penelitian ini adalah :

- Bagaimana mengembangkan dan implementasi model atau metode klasifikasi obyek suara yang beragam dari model Convolutional Neural Network (CNN) dan Support Vector Mechine (SVM).
- Melakukan penerapan metode yang digunakan dari klasifikasi obyek suara yang beragam.
- Mengetahui model yang diusulkan dapat memberikan hasil model baru yang diusulkan untuk diterapkan pada klasifikasi obyek suara yang beragam.
- Melakukan evaluasi serta pengembangan dari model CNN dan SVM dengan waktu komputasi yang lebih cepat dan akurasi yang lebih baik.

1.3 Kontribusi

Kontribusi keilmuan pada penelitian ini pada sisi teknologi adalah pemanfaatan model atau metode klasifikasi suara yang beragam dengan obyek yang diangkat pada kasus ini adalah suara burung dari model Convolutional Neural Network (CNN) dan Support Vector Mechine (SVM). Penelitian ini juga diharapkan dapat mengembangkan model yang dapat mendeteksi spesies burung berdasarkan suara burung di atas suara latar lainnya di lingkungan alam.

Kontribusi kelimuan pada bidang ornitologi dan ekologi yaitu menjadi pemanfaatan model pada alat yang sangat berguna dan memungkinkan untuk pemahaman yang lebih dalam mengenai kembang biak, evolusi, keanekaragaman hayati lokal dan bahkan perubahan iklim.

Bab 2

Tinjauan Pustaka

2.1 Suara Burung

Suara burung merupakan salah satu penanda khas adanya keberadaan burung. Populasi burung Kicauan burung yang cepat dan bervariasi mengakibatkan sinyal suara burung memiliki amplitudo yang bervariasi, sehingga sinyal suara burung merupakan sinyal non-stationer. (Miftahul Afida Annisa, 2020).

Suara burung bersifat tonal, dengan kehadiran komponen harmonik yang penting yang mencapai frekuensi tinggi (seringkali di atas 14 kHz). Suara serangga menunjukkan lebih banyak variabilitas. Beberapa dari mereka memusatkan energi pada frekuensi tertentu, sementara yang lain memiliki energi yang didistribusikan dalam pita yang berbeda, atau suara nada. (Diez Gaspon Itxasne, dkk, 2019).

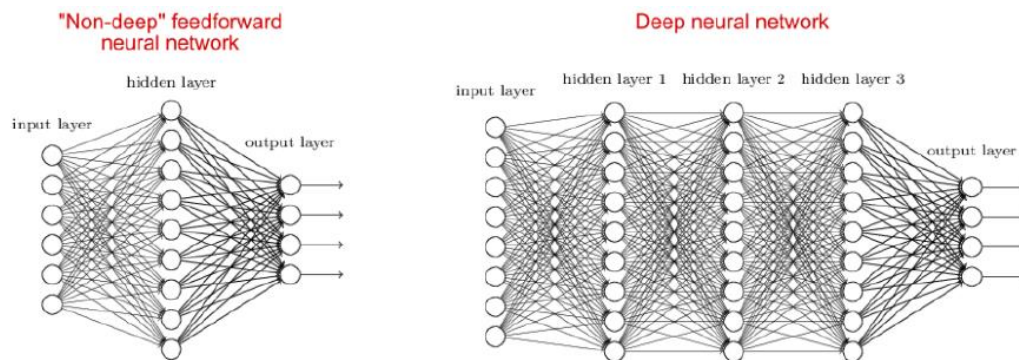
2.2 Deep Learning

Deep learning merupakan disiplin ilmu subset dari machine learning, tidak seperti teknik machine learning konvensional yang dibatasi oleh kemampuannya untuk memproses data mentah dan tergantung pada keahlian domain yang besar. Jika machine learning membutuhkan rekayasa yang cermat untuk merancang ekstraksi fitur, deep learning dapat mempelajari fitur dan memproses data secara langsung dalam bentuk mentah (LeCun, Y., Bengio, Y., Hinton, 2015).

Goodfellow et al (2016) mendefinisikan deep learning sebagai berikut: "Deep Learning adalah jenis machine learning yang memiliki kekuatan besar dan fleksibilitas yang tinggi dengan belajar untuk mewakili dunia sebagai hierarki konsep bersarang, dengan masing-masing konsep didefinisikan dalam kaitannya dengan konsep yang lebih sederhana, dan lebih banyak representasi abstrak yang dihitung dalam hal yang kurang abstrak" (Goodfellow et al., 2016).

Sebagai contoh jika ada tugas untuk mengklasifikasikan gambar yang diberikan, jika itu mewakili kucing atau anjing, teknik machine learning konvensional harus mendefinisikan fitur wajah seperti telinga, mata, kumis, mulut dan sebagainya,

maka perlu menulis metode untuk menentukan fitur mana yang lebih penting ketika mengklasifikasikan hewan tertentu, sedangkan deep learning tidak perlu menyediakan fitur secara manual, dengan deep learning fitur yang paling penting akan diekstraksi secara otomatis, setelah menentukan fitur mana yang paling penting untuk mengklasifikasi foto (Zaccone & Karim, 2018).



Gambar 1. Perbandingan Standar *Neural Network* dan *Deep Neural Network*

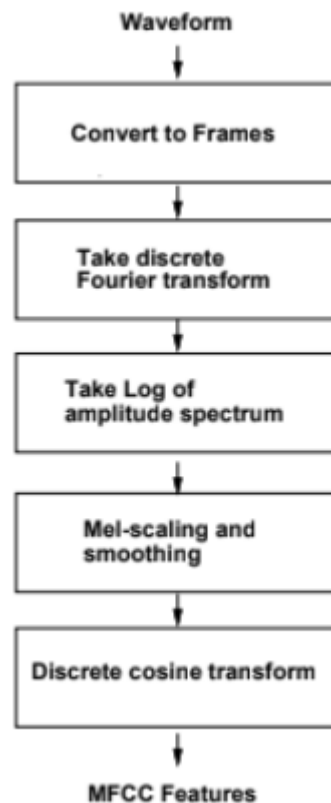
Implementasi dari *deep learning* biasanya menggunakan arsitektur *neural network* seperti yang ditunjukkan pada gambar 2.1, tetapi tidak seperti *neural network* tradisional yang biasanya hanya terdiri dari beberapa layer, *deep learning* biasanya terdiri dari ratusan bahkan ribuan layer untuk jaringan tersebut.

2.3 MFCC (Mel-Frequency Cepstral Coefficients)

MFCC (Mel Frequency Cepstral Coefficients) merupakan salah satu metode yang banyak digunakan dalam bidang speech technology, baik speaker recognition maupun speech recognition. Metode ini digunakan untuk melakukan feature extraction, sebuah proses yang mengkonversikan sinyal suara menjadi beberapa parameter. Spektrum suara menampilkan frekuensi berbeda yang ada dalam suara [7].

Mel Frequency Cepstrum(MFC) adalah representasi transformasi linear kosinus dari spektrum daya log waktu singkat dari sinyal ucapan pada skala frekuensi Mel non-linear. MFCC (*Mel-Frequency Cepstral Coefficient*) telah menjadi teknik yang sangat awam dan efisien untuk *signal processing*. MFCC mewakili model telinga, dan dapat menghasilkan pengenalan suara terutama ketika menggunakan jumlah koefisien yang tinggi. Gambar 2 menunjukkan proses dari pembuatan fitur MFCC.

Langkah pertama ada membagi sinyal menjadi banyak *frame*, biasanya menggunakan *windowing function* dengan interval tetap. Tujuannya adalah untuk memodelkan bagian kecil (biasanya 20 ms) dari sinyal yang diterima. *Window function*, atau disebut juga *Hamming window*, akan menghilangkan *edge effects*. *Cepstral feature vector* akan dibuat untuk setiap *frame*. (S Lalitha, D. G, 2015)



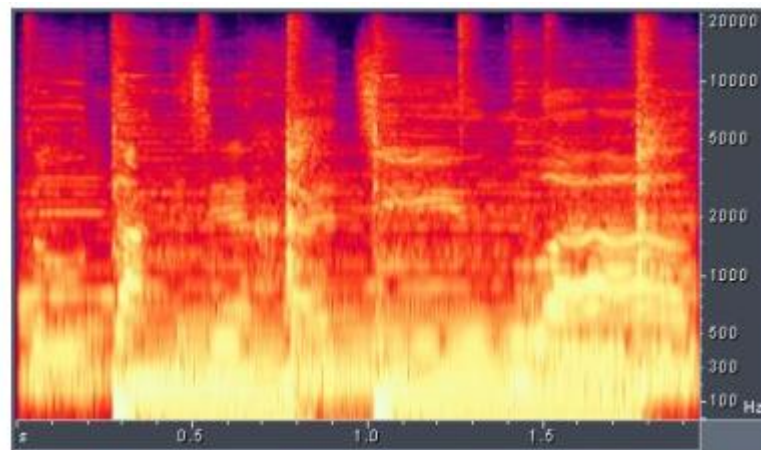
Gambar 2 – Proses untuk membuat fitur MFCC

2.4 Spektrogram

Gambar spektrogram adalah representasi visual dari spektrum frekuensi sinyal. Spektrogram adalah grafik yang menggambarkan perubahan frekuensi dan intensitas gelombang menurut sumbu waktu. Spektrogram digunakan dalam bidang musik, linguistik, sonar, radar, pengolahan wicara, seismologi, dan lain-lain. Spektrogram suara dapat digunakan untuk mengenali wicara secara fonetik dan untuk menganalisis ucapan hewan. Spektrogram dapat dibentuk dengan menggunakan *Fourier Transform*. Spektrogram didefinisikan sebagai besarnya kuadrat dari STFT, memberikan kekuatan suara untuk frekuensi dan waktu tertentu dalam dimensi ketiga.

Membuat spektrogram dengan FFT dilakukan dengan pertama mengambil data

yang berada pada *time domain*, dan memecah data-data tersebut menjadi beberapa bagian, dan dilakukan *Fourier Transform* untuk menghitung *magnitude* dari spektrum frekuensi untuk setiap bagian. Setelah itu tiap bagian akan disesuaikan pada garis vertical pada gambar, yang merupakan perbandingan *magnitude* dengan frekuensi dalam waktu tertentu. Setelah itu spektrum-spektrum tersebut akan di *plot* secara bersebelahan untuk membentuk suatu gambar. Contoh gambar spektrogram dapat dilihat pada gambar 3. (S Lalitha, D. G, 2015)

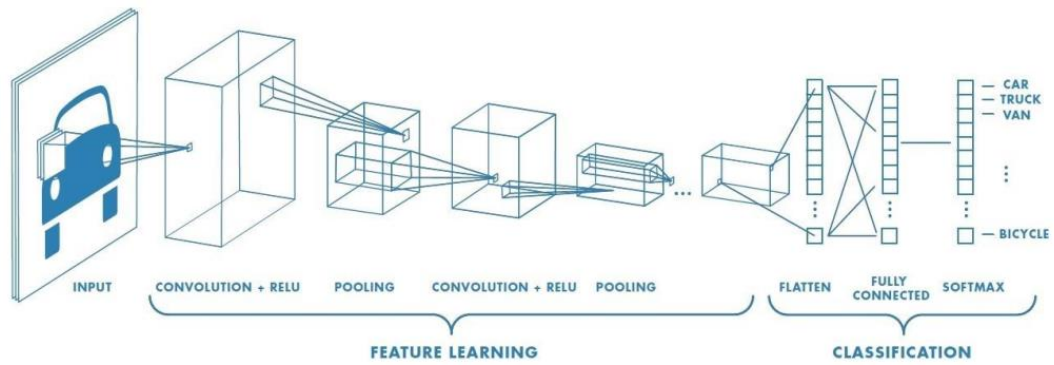


Gambar 3. Gambar Spektrogram

2.5 Arsitektur CNN

Convolutional Neural Network merupakan salah satu bagian dari *neural network* yang umumnya digunakan dalam pengolahan data *citra*. Nama *convolutional neural network* mengindikasikan bahwa jaringan tersebut menggunakan operasi matematika yang disebut konvolusi yang merupakan sebuah operasi linear. Dan merupakan jaringan saraf yang menggunakan konvolusi sebagai pengganti perkalian matriks umum, setidaknya satu dari *convolutional layer* tersebut.

Secara umum arsitektur dari *convolutional neural network* dibagi menjadi dua bagian besar, *Feature Extraction* atau *Feature Learning* dan *Classification*. Selain itu, arsitektur *Convolutional Neural Network* mempunyai tiga tipe lapisan utama yaitu: *convolutional layer*, *pooling layer* dan *fully connected layer*. *Fully-connected layer* yang terakhir disebut sebagai *output layer*. Variasi arsitektur *Convolutional Neural Network* untuk klasifikasi gambar/citra seperti Lenet, Alexnet, LeNet, AlexNet, VGG, GoogLeNet / Inception, ResNet dan banyak lagi. (Kyungdeuk Ko, dkk, 2018).



Gambar 4. Arsitektur CNN

2.6 Support Vector Machine (SVM)

SVM sebagai *machine learning* memperkenalkan metode baru yaitu untuk mencari *hyperplane* terbaik dari dua *class*. *Hyperplane* terbaik adalah *hyperplane* yang terletak ditengah-tengah antara dua set objek dari dua kelas yang ditentukan dengan mengukur *maximum margin* (maksimum margin) dari *hyperplane* dan mencari titik maksimalnya. Nilai margin menentukan jarak antar kelas, sedangkan *support vector* adalah titik terdekat dengan *hyperplane* atau tepat berada diatas *hyperplane* pada masing-masing kelas. Proses mencari lokasi *hyperplane* terbaik merupakan inti dari proses pembelajaran SVM. SVM dalam memroses suara digunakan untuk pembagian segmen fonetis, pengenalan suara, dan identifikasi audio secara general. (Miftahul Afida Annisa, 2020).

A. SVM Linier

SVM yang digunakan untuk menyelesaikan masalah linier yang disebut *Linear Separable Data* yaitu data yang dapat dipisahkan secara linear dari 2 *class* yaitu kelas -1 dan kelas +1 seperti pada gambar 2.1. Data yang masuk kelas -1 dirumuskan sebagai berikut:

$$(w \cdot x) + b = -1$$

Sedangkan untuk data yang masuk kelas +1 dirumuskan sebagai data yang memenuhi persamaan berikut:

$$(w \cdot x) + b = +1$$

Data dinotasikan sebagai $x_i \in \mathbb{R}^d$ sedangkan label masing-masing $y_i \in \{-1, +1\}$ untuk $i = 1, 2, 3, \dots, l$ dengan l adalah banyaknya data.

B. SVM Non-Linier

Terdapat kemungkinan bahwa titik pada dua kelas tidak dapat dipisah dengan sebuah *hyperplane* pada ruang yang sebenarnya. Untuk itu diperlukan ruang fitur yang lebih tinggi untuk mentransformasikan setiap data sehingga dapat dipisahkan secara linier oleh *hyperplane*.

Untuk mengatasi masalah tersebut, diperkenalkan suatu kernel sehingga pencarian *hyperplane* menjadi cepat. Selain itu, masalah yang tidak linear separable (nonlinier) dapat diatasi dengan fungsi kernel trick. Kernel trick merupakan fungsi yang memetakan fitur berdimensi rendah ke dimensi tinggi. Pada ruang vektor yang baru, *hyperplane* yang memisahkan kedua class dapat direkonstruksi.

Pemetaan dilakukan dengan tetap menjaga topologi suatu data yang berarti data yang berjarak dekat pada *input space* akan berjarak dekat pula pada *feature space*, begitu juga sebaliknya. Proses pembelajaran SVM untuk menemukan data sebagai titik support vector hanya bergantung pada *dot product* dari data yang sudah ditransformasikan ke ruang baru yang berdimensi tinggi yaitu:

$$\phi(\vec{x^i}) \cdot \phi(\vec{x^j})$$

Umumnya transformasi ϕ tidak diketahui, sehingga perhitungan *dot product* sesuai teori Mercer dapat digantikan fungsi kernel $K(\vec{x^i}, \vec{x^j})$ yang mendefinisikan secara implisit transformasi ϕ . Inilah yang disebut dengan Kernel trick:

$$K(\vec{x^i}, \vec{x^j}) = \phi(\vec{x^i}) \cdot \phi(\vec{x^j})$$

2.7 Perbandingan Tinjauan

Penelitian-penelitian terkait klasifikasi suara menggunakan model yang diajukan dan menjadi acuan penelitian disajikan pada tabel 1.

Tabel 1. Perbandingan Tinjauan Penelitian

Peneliti	Judul	Metode	Hasil Penelitian
Diez Gaspon Itxasnel, Saratxaga Ibon, Lopez de Ipiña Karmele [2019]	Deep Learning For Natural Sound Classification	Convolutional Neural Network	Akurasi keseluruhan sistem kami untuk klasifikasi 3 kategori tersebut adalah 79%, sedangkan akurasi yang diperoleh sistem klasifikasi suara lingkungan lainnya

			adalah 73%, 75% untuk 10 kategori dan 75,4% untuk 5 kategori.
J. Xie, K. Hu, M. Zhu, J. Yu, and Q. Zhu. [2019]	Investigation of Different CNN-Based Models for Improved Bird Sound Classification	Convolutional Neural Network	Hasil eksperimen pada klasifikasi 43 jenis burung menunjukkan bahwa menggabungkan model deep learning yang dipilih secara efektif dapat meningkatkan kinerja klasifikasi. Model fusi terbaik kami dapat mencapai akurasi seimbang 86,31% dan skor F1 tertimbang 93,31%.
Annisa Miftahul Afida [2020]	Klasifikasi Jenis Burung Berdasarkan Suara Menggunakan Algoritme Support Vector Machine	Support Vector Machine (SVM)	Hasil sinyal suara yang telah direduksi noise-nya menghasilkan sinyal suara yang lebih jernih dengan PSNR sebesar 4,2890. Klasifikasi dengan menggunakan SVM Kernel Heavy Tailed RBF dan rasio 7:3 menghasilkan akurasi sebesar 77,00% dengan parameter epsilon = $1.00e - 09$ dan $c = 10.000$.
Abdiel Willyar Goni, Deiby Tineke Salaki, dan Luther Alexander Latumakulita [2021]	Identifikasi Suara Burung Menggunakan Mel-Frequency Cepstral Coefficients Dan Backpropagation Neural Network	<i>Mel-Frequency Cepstral Coefficients</i> (MFCC) dan <i>Backpropagation Neural Network</i> (BPNN)	Sistem berhasil mengidentifikasi suara burung dengan <i>accuracy</i> tertinggi 95.83% dan <i>accuracy</i> terendah 93.33%.
Cecilia Tania Emanuella, Musfita, dan Armin Lawi [2021]	Klasifikasi Suara Kucing dan Anjing Menggunakan Convolutional Neural Network	Convolutional Neural Network	Hasil akurasi latih terbaik sebesar 93.8% dan akurasi validasi tertinggi 90%, untuk epoch terakhir akurasi latih 92.2% dan akurasi validasi 87.7% dengan <i>f1-score</i> 88%. Untuk <i>precision</i> kucing 87% dan anjing 89% serta <i>recall</i> kucing 89% dan anjing 87%, dimana label positifnya yaitu kucing dan skor AUROC 87.8%

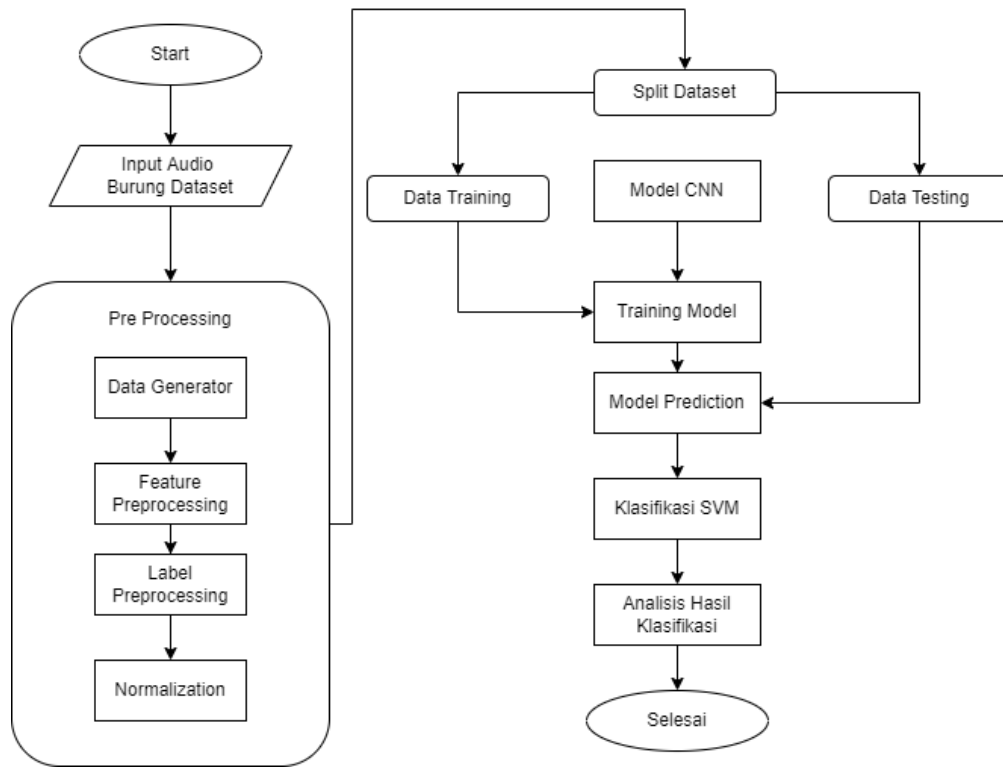
Bab 3

Metodologi

3.1 Gambaran Umum Penelitian

Metodologi penelitian yang digunakan dapat ditunjukkan pada gambar 5 dan dengan penjelasan sebagai berikut :

1. Pengumpulan data, pada tahapan ini data diambil dari data set terbuka seperti dataset Machine Learning for Signal Processing (MLSP) Bird Classification Dataset yang tersedia di bank data daring Kaggle. Biasanya, *Neural Network* menghasilkan keputusan klasifikasi untuk sebuah input agar bekerja sebagai sistem deteksi, sinyal yang ingin mendeteksi suatu peristiwa, harus dibagi menjadi fragmen-fragmen yang lebih kecil. Fragmen-fragmen ini digunakan sebagai input CNN sebagai klasifikasi.
2. Input Audio, pada tahapan ini dataset yang telah didapat kemudian dilakukan reduksi *noise* dengan tujuan mendapatkan sinyal suara yang lebih jernih.
3. Ekstraksi fitur dan pelabelan, pada tahap ini setiap data yang telah dibagi akan melalui proses ekstraksi fitur dengan menggunakan algoritma tertentu.
4. Melakukan pembagian data set untuk training dan testing yang berguna untuk mengenalkan pada program dengan model CNN.
5. Menyiapkan model CNN untuk proses training dan prediksi.
6. Melakukan pencatatan evaluasi dari performa pada model CNN.
7. Klasifikasi menggunakan algoritme SVM, pada tahapan ini data yang sudah melalui tahap prapemroses, kemudian diklasifikasikan menggunakan algoritme SVM. Output hasil klasifikasi adalah besar akurasi, waktu pelatihan dan pengujian
8. Melakukan analisis hasil pengujian dari klasifikasi yang didapatkan.



Gambar 5. Metodologi Penelitian

3.2 Rencana Kerja

Sebagai acuan rencana kerja pada penelitian ini, dapat dilihat pada tabel 2 dibawah ini.

Tabel 2 Rencana Pelaksanaan Penelitian

Kegiatan	Bulan Ke-																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Studi Pustaka																		
Pemilihan Topik																		
Penyusunan Proposal																		
Mencari dan mempelajari literatur																		
Mendesain model sesuai kebutuhan																		
Implementasi model yang diajukan																		
Pengujian Model																		
Publikasi Ilmiah/Seminar																		
Pengembangan atas saran seminar																		
Penyusunan Disertasi																		
Sidang Disertasi																		

Bibliografi

- A. Torfi, S. M. Iranmanesh, N. Nasrabadi, and J. Dawson, “3D convolutional neural networks for cross audio-visual matching recognition,” *IEEE Access*, vol. 5, pp. 22081–22091, 2017.
- Diez Gaspon Itxasne, Saratxaga Ibon, Lopez de Ipiña Karmele, “*Deep Learning For Natural Sound Classification*,” semantic scholar, Corpus ID: 204808728, Madrid, 2019.
- Elias Sprengel, Jaggi, M., Kilcher, Y., & Hofmann, T. (2016). Audio based bird species identification using deep learning techniques. CLEF.
- Emanuella Cecilia Tania, Musfita, Lawi Armin. “*Klasifikasi Suara Kucing dan Anjing Menggunakan Convolutional Neural Network*”, Konferensi Nasional Ilmu Komputer (KONIK), pp. 321-327. 2021
- Gibbens, S. (2018, September 5). These 8 Bird Species Have Disappeared This Decade. Retrieved from National Geographic: <https://www.nationalgeographic.com/environment/2018/09/news-macaw-extinct-bird-speciesdeforestation/>
- Goni Willyar Abdiel, Salaki Deiby Tineke, Latumakulita Luther Alexander, “*Identifikasi Suara Burung Menggunakan Mel-Frequency Cepstral Coefficients Dan Backpropagation Neural Network*,” Konferensi Nasional Ilmu Komputer (KONIK), pp. 9-13. 2021
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. In MIT Press. MIT Press. <https://www.deeplearningbook.org/>
- In H. S. Alikodra, *Pengelolaan Satwa Liar*. Bogor: Yayasan Penerbit Fakultas Kehutanan. (2002).
- J. Xie, K. Hu, M. Zhu, J. Yu, and Q. Zhu, “*Investigation of Different CNN-Based Models for Improved Bird Sound Classification*,” *IEEE Access*, vol. 7, pp. 175353–175361, 2019.
- Kyungdeuk Ko, Sangwook Park, Hanseok Ko, “*Convolutional Feature Vectors and Support Vector Machine for Animal Sound Classification*,” *IEEE Access*, vol. 5, pp. 376–379, 2018.
- LeCun, Y., Bengio, Y., Hinton, G. (2015). Deep learning. *nature* 521 (7553): 436. *Nature*, 521, 436–444.
- Miftahul Afida Annisa, “*Klasifikasi Jenis Burung Berdasarkan Suara Menggunakan Algoritme Support Vector Machine*,” *Jurnal Ilmiah Matematika*, vol. 8, pp. 1–6, 2020.

- N. Singh, “*Classification of Animal Sound Using Convolutional Neural Classification of Animal Sound Using Convolutional Neural Network Network*,” 2020.
- Raghuram, M. A., Chavan, N. R., Belur, R., & Koolagudi, S. G. (2016). Bird classification based on their sound patterns. Springer, 791-804.
- S Lalitha, D. G. 2015. Emotion Detection using MFCC and Cepstrum Features. *Procedia Computer Science* 70, 29-35.
- Upadhyay, R. K., Rai, V., & Iyengar, S. (2001). Species Extinction Problem: Genetic vs Ecological Factors. *Applied Mathematical Modelling* 25, 937-951.
- Zaccone, G., & Karim, M. R. (2018). *Deep Learning with TensorFlow: Explore neural networks and build intelligent systems with Python (2nd ed.)*. Packt Publishing.