

Klasifikasi Suara Burung Memanfaatkan *Bidirectional Long Short Term Memory*

Gregorio Hartono¹, Hendry Setiawan², Kestriia Rega Priianti³

Program Studi Teknik Informatika, Universitas Ma Chung
Villa Puncak Tidar Blok N no. 1, Doro, Karangwidoro, Kec. Dau, Kabupaten Malang, Jawa Timur, Indonesia

¹311910008@student.machung.ac.id

²hendry.setiawan@machung.ac.id

³kestriia.rega@machung.ac.id

Dikirim pada 22-11-2024, Direvisi pada 27-11-2024, Diterima pada 04-12-2024

Abstrak

Indonesia memiliki berbagai macam jenis burung dilindungi tersebar luas dan harus dilestarikan. Pelestarian burung menggunakan suara menjadi salah satu cara untuk membedakan jenis burung yang dilindungi. Oleh sebab itu dalam penelitian ini melibatkan dataset 7 suara burung dilindungi dalam bentuk mp3 dan dilakukan preprocessing yaitu *noise reduction* dan pemotongan suara burung untuk pelatihan data. Selanjutnya pembuatan model *deep learning* menggunakan arsitektur *Bidirectional Long Short Term Memory* (BiLSTM). Total data untuk penelitian ada 560 suara berdurasi 4 detik yang dalam pembuatan modelnya dibagi menjadi 64% data latih model, 16% untuk data validasi, dan 20% untuk data uji. Pencarian *hyperparameter* dari 360 percobaan didapatkan pada ReLU dan Adam dengan rata rata nilai keseluruhan 98,579% dalam mengklasifikasikan suara burung yang dilindungi.

Kata Kunci: Bidirectional Long Short-Term Memory, Hyperparameter, Klasifikasi suara burung, Recurrent Neural Network, Suara Burung.

Ini adalah artikel akses terbuka di bawah lisensi [CC BY-SA](#).



Penulis Koresponden:

Gregorio Hartono
Program Studi Teknik Informatika, Universitas Ma Chung, Villa Puncak Tidar Blok N no. 1, Doro, Karangwidoro, Kec. Dau, Kabupaten Malang, Jawa Timur 65151 Indonesia Email: 311910008@student.machung.ac.id

I. PENDAHULUAN

Keanekaragaman hayati Indonesia sangat beragam. Keanekaragaman hayati harus senantiasa dilestarikan dan dilindungi, dan pemerintah dan masyarakat bertanggung jawab untuk melakukannya. Perlindungan dan pelestarian sumber daya alam Indonesia terkait dengan keanekaragaman hayati. Sumber daya alam tersebut termasuk penggolongan, lingkungan hidup, dan ekosistemnya, penggolongannya termasuk sumber daya alam hayati dan non hayati. Tujuan konservasi adalah untuk melindungi dan melestarikan keanekaragaman hayati yang hampir punah atau langka di Indonesia. Saat ini, banyak tumbuhan dan satwa langka khas Indonesia perlu dilestarikan karena populasinya semakin berkurang. Dengan konservasi, manusia diharapkan untuk menggunakan sumber daya alam dengan cermat [1]. Konservasi menurut *World Conservation Strategy* merupakan kegiatan manajemen biosfer oleh manusia dengan tujuan dapat dipelihara secara berkelanjutan untuk generasi mendatang[2].

Namun, upaya konservasi ini terbalik dengan tindakan eksploitasi oleh pihak yang tidak bertanggung jawab terutama pada area hutan. Eksploitasi menurut KBBi adalah sebuah tindakan pemanfaatan yang dilakukan demi keuntungan pribadi yang tidak dapat dibenarkan[3]. Salah satu tindakan eksploitasi terhadap satwa yang dilindungi adalah perdagangan ilegal. Praktik perdagangan ilegal meliputi proses perburuan, pengangkutan, penyiwaan/pembunuhan, pengiriman, pemindahtanganan, penampungan, dan penerimaan

satwa dengan tujuan eksploitasi[4]. Perawatan di konservasi terhadap satwa yang dilindungi harus dilakukan dengan hati-hati oleh sebab itu klasifikasi menggunakan suara menjadi salah satu solusi yang dapat dilakukan. Sexing dengan suara cukup mudah dilakukan dengan cara merekam suara burung tanpa menyentuh dan meraba-raba badannya untuk mengecek jenis kelamin. Suara yang dikeluarkan berbeda beda tergantung kondisinya seperti ketika memanggil anak, menarik perhatian lawan jenis, pertahanan wilayah, dan berkomunikasi satu dengan yang lainnya. Variasi suara setiap burung bisa berbeda-beda dapat menjadi penentu suara yang berasal dari burung yang dilindungi dari jenis maupun ras tertentu.

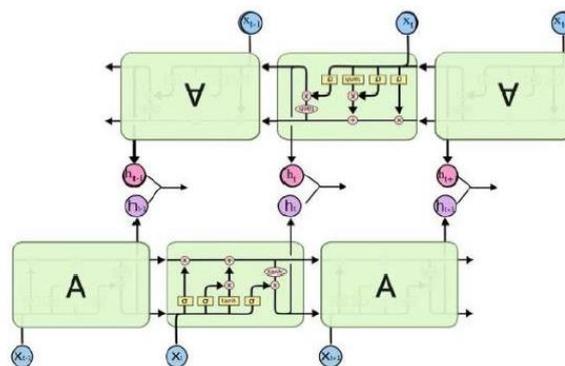
Sebelumnya telah dilakukan penelitian oleh Funay pada tahun 2021 dalam tugas akhirnya berjudul Rancang Bangun Aplikasi Pengenalan Suara Burung dengan *Long Short Term Memory* (LSTM). Metode yang digunakan berasal dari *Recurrent Neural Network* (RNN) dengan model LSTM dihasilkan 3 persentase prediksi dari 2 uji coba yang berbeda. Pertama menggunakan aplikasi *android* menghasilkan 74,29% dan 40,48% ,sedangkan kedua dengan Python dan *web client* menghasilkan 57,14%, sehingga dari hasil prediksi tersebut masih memiliki potensi untuk ditingkatkan [5]. Penelitian yang dilakukan oleh Liu pada tahun 2021 dalam proses pembuatan model dilakukan proses klasifikasi, filter, dan ekstrak fitur *Mel Frequency* dalam *TFRRecord* dilanjutkan dengan membangun arsitektur BiLSTM dan dilakukan uji coba dengan mengklasifikasi 20 spesies burung menggunakan softmax dan sebanyak 14.311 file audio didapatkan hasil rata-rata 92,2% yang membuat BiLSTM lebih baik dibandingkan LSTM [6].Oleh sebab itu, penelitian ini akan melakukan uji dengan *Bidirectional Long Short Term Memory* (BLTSM) dengan tujuan dapat menguji arsitektur BiLSTM dalam mengenali suara burung dan melakukan klasifikasi yang sesuai.

Masalah yang dihadapi pada penelitian ini adalah melakukan identifikasi terhadap jenis burung dilindungi pada endemik Indonesia. Beberapa pasar burung di Indonesia masih tercatat melakukan aktivitas perdagangan satwa dilindungi seperti Pasar Burung Sukahaji di Bandung, Pasar Burung Pramuka di Jakarta Timur, Pasar Burung Jatinegara di Jakarta Timur, Pasar Pal 7 di Banjarmasin, dan Pasar Burung Plered di Cirebon[7]. Hal ini menjadi salah satu ancaman serius terhadap konservasi satwa di Indonesia. Dalam undang-undang No.5 tahun 1990 tentang Konservasi Sumber Daya Alam Hayati dan Ekosistemnya dalam pasal 40 ayat (2) menjelaskan bahwa segala tindakan yang mengeksploitasi satwa yang dilindungi akan dikenai sanksi pidana[8][9].

Dari permasalahan yang dipaparkan sebelumnya, penelitian yang akan dilakukan bertujuan untuk mengembangkan model arsitektur BiLSTM untuk proses klasifikasi suara dari burung yang dilindungi pada endemik Indonesia, sehingga dapat memberikan peluang orang yang tidak terlalu mengenal burung dapat mengidentifikasi burung yang dilindungi melalui suara.

II. METODE PENELITIAN

Bidirectional Long Short Term Memory (BiLSTM) merupakan varian dari *Long Short Term Memory* (LSTM) yang merupakan sebuah tipe arsitektur RNN. LSTM memiliki kapasitas besar untuk menyimpan informasi (memori jangka panjang) dan menjaga relevansi keadaan terkini (memori jangka pendek) dengan struktur yang hanya berfokus terhadap satu arah saja. BiLSTM adalah dua layer LSTM yang masing masing memiliki arah yang berbeda seperti pada Gambar 1. Layer LSTM pertama digunakan untuk keadaan masa lalu dan LSTM kedua untuk keadaan masa depan [10]. Sehingga dari hal tersebut bisa dipahami bahwa perbedaan LSTM dan BiLSTM terdapat pada inputan. LSTM hanya memiliki inputan dari satu arah saja, sedangkan BiLSTM memiliki inputan dari dua arah sehingga inputan dan output bisa berasal dari dua layer LSTM yang dimiliki Bi-LTSM.



Gambar 1. Model Bi-LSTM [11]

$$y_t = W \rightarrow_{hy} \vec{h}_t + W \leftarrow_{hy} \overleftarrow{h}_t \quad (1)$$

Keterangan:

y_t = output gate LSTM dua arah

$W \rightarrow_{hy}$ = nilai boot untuk *output gate* LSTM maju

\vec{h}_t = nilai *output* LSTM maju

$W \leftarrow_{hy}$ = nilai boot untuk *output gate* LSTM mundur

\overleftarrow{h}_t = nilai *output* LSTM mundur

Pada Gambar 1. terdapat *hidden layer* yang terhubung dari LSTM maju dan LSTM mundur. Dengan adanya *hidden layer* dua arah yang saling berlawanan maka model dapat memahami data dari depan dan belakang, sehingga proses pelatihan data lebih memahami data pada *time series*.

Fungsi aktivasi terjadi pada *output* setiap *neuron* dalam jaringan saraf tiruan. Fungsi ini memperkenalkan non-linearitas ke dalam jaringan, memungkinkan jaringan mempelajari hubungan kompleks antara *input* dan *output*. Pada arsitektur *deep learning* terdapat beberapa fungsi aktivasi yang digunakan yaitu :

a) *Tanh*

Fungsi aktivasi tanh atau *hyperbolic tangent* diperkenalkan oleh McCulloch dan Pitts pada tahun 1943. Tanh merupakan fungsi matematis yang sering digunakan sebagai fungsi aktivasi di dalam jaringan saraf. Penggunaan umum fungsi tanh adalah di dalam lapisan tersembunyi (*hidden layer*) dari jaringan saraf, terutama ketika data yang digunakan bersifat simetris dan memiliki nilai yang dapat berkisar positif dan negatif [12]. Fungsi tanh mengonversi nilai input ke rentang antara -1 dan 1. Fungsi tanh didefinisikan sebagai:

$$\tanh(x) = \frac{(e^x - e^{-x})}{(e^x + e^{-x})} \quad (2)$$

Keterangan:

e = nilai euler (sekitar 2.71828)

x = nilai input

b) *ReLU*

ReLU, atau *Rectified Linear Unit*, adalah salah satu jenis fungsi aktivasi yang umum digunakan dalam jaringan saraf tiruan (*neural networks*). Fungsi ini banyak digunakan karena sederhana namun efektif dalam memecahkan masalah *vanishing gradient* yang dapat terjadi pada fungsi aktivasi lainnya, seperti sigmoid atau tanh. Fungsi sigmoid didefinisikan sebagai:

$$ReLU(x) = \max(0, x) \quad (3)$$

Keterangan:

x = nilai input

Artinya, jika nilai input (x) lebih kecil dari 0, maka fungsi akan menghasilkan nilai 0; sedangkan jika nilai input lebih besar dari atau sama dengan 0, maka fungsi akan menghasilkan nilai input tersebut. Dengan kata lain, ReLU hanya memasukkan nilai positif ke depan, sementara nilai negatif dibuat menjadi 0.

d) *Sigmoid*

Sigmoid adalah salah satu fungsi aktivasi yang digunakan dalam jaringan saraf tiruan. Fungsi sigmoid memiliki bentuk kurva S dan digunakan untuk mengonversi input berkelanjutan menjadi output dalam rentang antara 0 dan 1. Fungsi sigmoid didefinisikan sebagai berikut:

$$Sigmoid(x) = \frac{1}{1 + e^{-x}} \quad (4)$$

Keterangan:
 x = nilai input

Fungsi sigmoid memiliki bentuk kurva S atau "S-shaped curve", dan karakteristik ini membuatnya berguna dalam banyak konteks, terutama dalam pembuatan prediksi probabilitas.

e) *Softmax*

Softmax adalah fungsi matematis yang umum digunakan dalam jaringan saraf tiruan, khususnya pada lapisan output untuk tugas klasifikasi multikelas. Fungsi *softmax* mengonversi nilai input menjadi distribusi probabilitas yang mendekati distribusi diskrit, di mana probabilitas dari setiap kelas diukur sebagai eksponensial dari nilai input dibagi oleh jumlah eksponensial dari semua nilai input.

$$\text{Softmax}(x)_i = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}} \quad (5)$$

Keterangan:

e = nilai euler (sekitar 2.71828)

$(x)_i$ = nilai input untuk kelas ke- i dalam vektor x

K = jumlah total kelas

i = indeks kelas tertentu yang sedang dievaluasi dalam konteks distribusi probabilitas *softmax*

$\sum_{j=1}^K e^{x_j}$ = jumlah dari semua nilai eksponensial dalam vektor x , dan merupakan faktor normalisasi yang membuat hasil *softmax* menjadi distribusi probabilitas.

Fungsi optimasi dalam *deep learning* merujuk pada algoritma yang digunakan untuk mengoptimalkan atau meminimalkan fungsi kerugian (*loss function*) selama pelatihan model. Tujuannya adalah untuk menemukan parameter model yang menghasilkan prediksi yang paling akurat dengan meminimalkan kesalahan prediksi.

f) *Adam*

Adaptive Moment Estimation atau disingkat Adam merupakan kombinasi dari metode RMSprop dan Momentum. Adam mengadaptasi laju pembelajaran untuk setiap parameter secara dinamis berdasarkan perkiraan momen pertama dan kedua dari gradien.

g) *RMSProp*

Root Mean Square Propagation atau disingkat RMSprop adalah Algoritma optimasi yang menyesuaikan laju pembelajaran untuk setiap parameter berdasarkan perkiraan dari perubahan rata-rata kuadrat

h) *Nadam*

Nesterov-accelerated Adaptive Moment Estimation atau disingkat Nadam merupakan algoritma optimasi yang menggabungkan dua konsep utama: *Nesterov Accelerated Gradient* (NAG) dan *Adaptive Moment Estimation* (Adam). Nadam dirancang untuk memperbaiki beberapa kelemahan yang terdapat pada Adam, terutama ketika berurusan dengan masalah yang memiliki permukaan loss yang melengkung

i) *SGD*

Stochastic Gradient Descent atau disingkat SGD merupakan algoritma optimasi yang paling dasar dan umum digunakan dalam pelatihan model pada *deep learning*. SGD termasuk dalam kategori optimasi berbasis gradien yang bekerja dengan meminimalkan fungsi kerugian (*loss function*) untuk memperbarui parameter model.

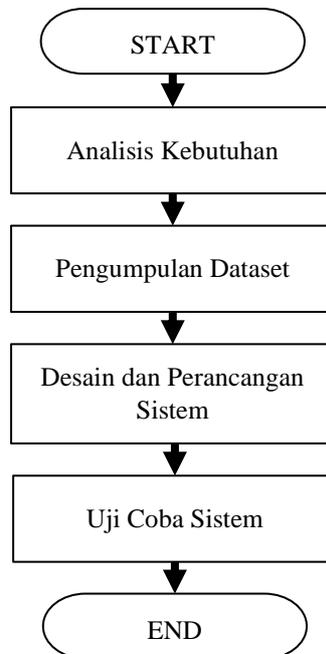
j) *Adadelta*

Adadelta adalah algoritma optimasi yang dikembangkan sebagai pengembangan dari algoritma Adagrad. Algoritma ini dimaksudkan untuk mengatasi beberapa masalah Adagrad, terutama terkait dengan penurunan laju pembelajaran yang terlalu cepat seiring waktu.

Mel-Frequency Cepstral Coefficients (MFCC) adalah fitur audio dalam pengenalan suara dan pemrosesan sinyal. MFCC menggunakan skala frekuensi *spasi quasi-logarithmic*, yang mirip dengan cara kerja pendengaran manusia memproses suara. Kelebihan MFCC adalah mampu menangkap karakteristik suara yang terkandung dalam sinyal suara, menghasilkan data seminimal mungkin tanpa menghilangkan informasi-informasi penting yang ada, dan mengadaptasi organ pendengaran manusia dalam melakukan persepsi terhadap sinyal suara. Proses MFCC terdiri dari *pre-emphasis*, *frame blocking*, *windowing*, *Fast Fourier Transform* (FFT), *Mel Frequency Wrapping* (MFW), dan *Discrete Cosine Transform* (DCT) [13]

Pre-emphasis merupakan proses penyaringan yang bertujuan untuk mendapatkan bentuk spektral dari frekuensi sinyal suara yang lebih bagus. *Frame blocking* digunakan untuk meminimalisir sinyal yang terganggu atau hilang (*deletion*). Proses ini akan terus berlanjut hingga seluruh sinyal yang ada telah masuk ke dalam frame. *Windowing* berfungsi untuk meredam *noise* yang muncul di kedua ujung *frame*. *Fast Fourier Transform* (FFT) berfungsi untuk Merubah domain waktu sinyal menjadi domain frekuensi. *Mel Frequency Wrapping* (MFW) berfungsi untuk filter berupa *filterbank* untuk mengetahui ukuran energi dari *frequency band* tertentu dalam signal suara. *Discrete Cosine Transform* (DCT) berfungsi untuk mengembalikan sinyal suara pada domain frekuensi ke domain waktu, sehingga didapatkan koefisien cepstrum.

Penelitian ini dilakukan dengan melakukan uji coba terhadap arsitektur BiLSTM untuk dapat mengenali suara dari 7 jenis burung dengan beberapa perbandingan hyperparameter seperti jumlah layer dari BiLSTM, durasi audio, nilai *dropout*, fungsi aktivasi, dan fungsi optimasi. Proses Penelitian ini dibagi dalam beberapa tahap seperti pada Gambar 2.



Gambar 2. Tahapan Proses Penelitian

Tahapan alur penelitian yang pertama analisis kebutuhan, yaitu mempersiapkan segala kebutuhan yang terkait dengan proses penelitian. Tahap kedua, pengumpulan dataset yang merupakan data audio dari tujuh jenis burung dilindungi pada endemik Indonesia yang digunakan sebagai data pelatihan pada model sistem. Tahap ketiga, desain dan perancangan sistem yang dilakukan setelah mendapatkan model sistem yang mendukung proses klasifikasi suara burung dan dilanjutkan dengan uji coba klasifikasi suara burung untuk mengecek akurasi yang dimiliki oleh sistem yang telah dibuat.

1. Analisis Kebutuhan

Dalam penelitian ini dibutuhkan analisis sebagai persiapan agar penelitian dapat dilaksanakan dengan lancar. Kebutuhan dari sisi peneliti *Software* untuk mengolah bahasa pemrograman Python yaitu Visual studio code dengan beberapa *library* seperti Librosa, TensorFlow, Keras, Matplotlib, JSON, Math, Scikit-Learn/ Sklearn.

2. Pengumpulan Data

Pada saat ini, data dikumpulkan dari ketujuh jenis burung yang dilindungi oleh pemerintah: Jalak Bali (*Leucopsar rothschildi*), Nuri Tanau (*Psittinus cyanurus*), Kacamata/Pleci Jawa (*Zosterops flavus*), Opor Jawa (*Heleia javanica*), Gelatik Jawa (*Lonchura oryzivora*), Nuri Ternate (*Lorius garrulus*), Nuri Kepala Hitam (*Lorius lory*) dan Jalak Bali (*Leucopsar rothschildi*). Data suara dari ketujuh jenis burung dikumpulkan dari *website* dan *youtube*. Kriteria untuk memilih sumber untuk mengunduh suara burung ini adalah bahwa sumber tersebut hanya mengandung suara dari satu jenis burung dan tidak mengandung suara burung lain di background. Data suara dalam bentuk MP3 diambil dari *website* seperti *xeno-canto.org*, *e-bord.org*, *omkicau.com* dan *youtube* tersebut. Sebelum data dapat dilatih, tahap preprocessing akan dilakukan.

3. Desain dan Perancangan Sistem

Pada tahapan ini dilakukan desain dan perancangan sistem dari penelitian ini. Gambar 3 menunjukkan tahapan mengumpulkan dataset suara burung dengan format file mp3, kemudian file mp3 akan diubah menjadi WAV, dilanjutkan pengurangan *noise* dan pemotongan dataset dari audio burung ketika berbunyi, lalu audio hasil pengeditan akan diberikan label suara, dilanjutkan dengan *preprocessing* menggunakan MFCC, pelatihan model, kemudian dilakukan uji coba untuk menguji akurasi model untuk mengecek kecocokannya dengan dataset yang dimiliki.



Gambar 3. Tahap *Preprocessing*, *training*, dan Uji Coba model BiLSTM

a) Preprocessing

Pada proses ini, suara burung diunduh dalam bentuk MP3 dari *website* seperti *xeno-canto.org*, *e-bord.org*, *omkicau.com* dan *youtube*. Jumlah file suara yang dikumpulkan oleh setiap burung berbeda-beda, dan durasi file juga berbeda-beda. Burung Gelatik Jawa memiliki enam file suara, dengan durasi paling pendek 17 detik dan paling panjang 10 menit. Dari enam file tersebut, terdapat sembilan variasi suara burung Gelatik Jawa, empat file berisi suara burung Nuri Kepala Hitam dengan durasi paling sedikit 1 menit dan paling panjang 22 menit, dan empat file berisi suara burung Nuri Tanau dengan durasi paling sedikit 31 detik dan paling panjang 10 menit. Dari empat file tersebut, ada enam variasi suara burung Nuri Tanau; dua file dengan durasi paling sedikit 1 menit dan paling panjang 6 menit; dua file dengan 10 variasi suara burung Nuri Ternate; dua file dengan suara burung Opor Jawa dengan durasi paling sedikit 10 menit dan paling panjang 41 menit; dan dua file dengan suara burung Jalak dengan durasi paling sedikit 10 menit dan paling panjang 41 menit. Semua variasi suara burung Jalak terdiri dari sepuluh file, terdapat sepuluh variasi suara burung Pleci. Setiap file memiliki durasi 12 detik dan paling lama 11 menit.

Kemudian format suara diubah menjadi WAV dengan menggunakan *software Audacity* dan dipotong menjadi sepanjang 4 detik dengan tujuan untuk melakukan uji coba identifikasi yang berfokus pada suara setiap burung yang terbagi menjadi 1 detik, 2 detik, 3 detik, dan 4 detik untuk membandingkan tingkat akurasi klasifikasi. Sehingga didapatkan 80 suara untuk masing-masing kelas dan total untuk keseluruhan kelas adalah 560 suara. Selanjutnya, data tersebut akan dibagi ke dalam tiga bagian, yaitu data latih, data validasi, dan data uji. Dari jumlah 560 suara tersebut, 64% digunakan sebagai data latih model LSTM, 16% digunakan sebagai data validasi, dan 20% digunakan sebagai data uji. Jumlah data suara yang digunakan sebagai data latih adalah 358 suara, 90 suara digunakan sebagai data validasi, dan 112 suara digunakan sebagai data uji.

b) Training Model

Pada tahap ini akan melatih model dengan menggunakan RNN. Input *shape* yang diberikan berupa input dua dimensi. Desain RNN yang digunakan adalah menggunakan arsitektur BiLSTM. *Layer* yang digunakan untuk arsitektur BiLSTM akan diuji coba dalam 3 lapisan yaitu 64 *layer*, 128 *layer*, dan 256 *layer*. Durasi yang digunakan untuk audio terbagi menjadi 4 yaitu 1 detik, 2 detik, 3 detik, dan 4 detik. Fungsi aktivasi yang digunakan adalah Tanh, ReLU, dan Sigmoid, sedangkan untuk fungsi aktivasi output menggunakan *softmax*. Metode optimasi yang digunakan adalah Adam, RMSProp, Nadam, SGD, dan Adadelta. Seluruh metode optimasi menggunakan *learning rate* sebesar 1.00×10^{-4} atau 0.0001. *Dropout* yang digunakan ada 2 yaitu 0.3 dan 0.5. *Loss function* yang digunakan adalah *sparse categorical cross entropy*, *metrics* yang digunakan adalah *accuracy*, *batch size* yang digunakan adalah 32. Jumlah epoch yang digunakan sebanyak 500 epoch, dan menggunakan *early stopping* dengan *monitor loss* dan *patience* sebesar 3. Dari hasil percobaan tersebut akan dipilih satu model terbaik untuk selanjutnya disimpan dalam format h5 (ekstensi .h5). Selanjutnya model tersebut akan dilakukan uji coba dengan dataset yang digunakan dalam pembuatan model untuk menguji akurasi dan akhirnya dibandingkan antara metode Bi-LSTM.

c) Persiapan Uji Coba

Untuk mempermudah peneliti dalam melakukan uji coba dengan berbagai macam variasi yang telah dijelaskan pada poin sebelumnya, maka peneliti membuat tabel kelompok percobaan yang terdiri dari 24 kelompok dengan perbedaan dari durasi yang digunakan, *dropout*, dan *layer* dari Bi-LSTM yang ditunjukkan pada table 1.

Table 1. Tabel kelompok untuk mempermudah proses uji coba

No	Kelompok Percobaan	<i>Duration</i>	<i>Dropout</i>	<i>Layer</i>
1	A	2	0,3	128
2	B	3	0,3	128
3	C	4	0,3	128
4	D	2	0,5	128
5	E	3	0,5	128
6	F	4	0,5	128
7	G	2	0,3	64
8	H	3	0,3	64
9	I	4	0,3	64
10	J	2	0,5	64
11	K	3	0,5	64
12	L	4	0,5	64
13	M	2	0,3	256
14	N	3	0,3	256
15	O	4	0,3	256
16	P	2	0,5	256
17	Q	3	0,5	256
18	R	4	0,5	256
19	S	1	0,3	128
20	T	1	0,5	128
21	U	1	0,3	64
22	V	1	0,5	64
23	W	1	0,3	256
24	X	1	0,5	256

Setiap kelompok akan terdiri dari 15 percobaan dari kombinasi fungsi aktivasi Tanh, ReLU, dan Sigmoid dengan metode optimasi Adam, RMSProp, Nadam, SGD, dan Adadelta. Sehingga dari table 1. yang merupakan gambaran dari seluruh uji coba yang dilakukan didapatkan 360 percobaan untuk arsitektur BiLSTM

III. HASIL DAN PEMBAHASAN

Setelah dilakukan pembuatan model h5, akan dilakukan pengujian terhadap model dengan pencocokan terhadap dataset yang digunakan untuk membuat model. Hal ini bertujuan untuk mengecek hasil akurasi prediksi dari model yang telah dibuat terhadap suara dari jenis burung yang ada pada dataset dengan menggunakan *hyperparameter* yang telah ditetapkan pada setiap kelompok pada poin sebelumnya dalam dan Bi-LSTM.

```
DATASET_PATH = r"D:\1a.PKL\PKL_Audio"
JSON_PATH = "data2s.json"
SAMPLE_RATE = 22050
DURATION = 2 # measured in seconds
SAMPLES_PER_TRACK = SAMPLE_RATE * DURATION
```

Gambar 4. Parameter untuk mengubah durasi audio yang digunakan contohnya 2 detik

Pada Gambar 4. digunakan untuk melakukan pengujian parameter terhadap durasi yang akan digunakan untuk panjang audio dari suara burung yang akan digunakan, parameter yang digunakan adalah 1 detik, 2 detik, 3 detik, dan 4 detik.

```
model.add(keras.layers.Bidirectional(keras.layers.LSTM(128,
return_sequences=True), input_shape=input_shape))
model.add(keras.layers.Bidirectional(keras.layers.LSTM(128)))
```

Gambar 5. Parameter untuk mengubah layer yang digunakan BiLSTM contohnya 128 layer

Pada Gambar 5. digunakan untuk melakukan pengujian parameter terhadap layer BiLSTM yang akan digunakan membuat model BiLSTM, parameter yang digunakan adalah 64 layer, 128 layer, dan 256 layer.

```
model.add(keras.layers.Dense(32, activation='tanh'))
model.add(keras.layers.Dropout(0.3))
optimiser = keras.optimizers.Adam(learning_rate=0.0001)
```

Gambar 6. Parameter fungsi optimasi, *droup out*, dan metode optimiser

Pada Gambar 6. digunakan untuk melakukan pengujian parameter terhadap fungsi optimasi yaitu Tanh, ReLU, dan Sigmoid. Dilanjutkan dengan parameter *Dropout* yaitu 0.3 dan 0.5 , dan terakhir parameter metode optimasi yaitu Adam, RMSProp, Nadam, SGD, dan Adadelta.

Dari 360 percobaan yang telah dilakukan didapatkan hasil percobaan untuk melakukan prediksi atau *predicted class* untuk mengecek keberhasilan model yang telah dibuat untuk mengklasifikasi jenis suara burung yang dimiliki. Hasil uji coba dalam penelitian ini dapat dilihat pada table 2. yang menunjukkan 360 hasil percobaan terhadap kecocokan prediksi model untuk mengklasifikasikan suara burung. Pada Table 2. Terdapat area yang berwarna kuning. Jika dilihat berdasarkan fungsi aktivasi dan optimasi model dengan fungsi ReLU dan Adam memiliki rata-rata nilai paling tinggi dibandingkan dengan fungsi lainnya yaitu sebesar 98,579%, sedangkan jika dilihat durasi, *dropout*, dan *layer* didapatkan rata-rata terbaik pada percobaan R dengan durasi 4detik, dropout 0.5, dan *layer* 256. Percobaan R memiliki nilai rata-rata 71,393%

Table 2. Hasil uji coba model terhadap data suara burung sebanyak 560 suara

Perco baan	SIGMOID					RELU					TANH				
	ADADE LTA	SG D	NAD AM	RMSD ROP	AD AM	ADADE LTA	SG D	NAD AM	RMSD ROP	AD AM	ADADE LTA	SG D	NAD AM	RMSD ROP	AD AM
A	13,750%	13,75 0%	98,571 %	98,036 %	99,10 7%	13,214%	20,71 4%	99,464 %	98,571 %	99,28 6%	16,964%	23,92 9%	99,107 %	98,214 %	98,75 0%
B	14,286%	14,28 6%	99,286 %	99,107 %	95,17 9%	15,000%	17,85 7%	99,643 %	98,571 %	98,75 0%	14,643%	20,89 3%	99,107 %	99,286 %	98,92 9%
C	20,536%	9,821 %	99,643 %	99,286 %	98,03 6%	11,250%	22,67 9%	99,643 %	99,286 %	98,03 6%	9,464%	30,00 0%	95,714 %	99,643 %	99,64 3%
D	14,107%	14,28 6%	99,464 %	98,571 %	98,75 0%	17,857%	15,53 6%	99,286 %	98,214 %	98,39 3%	17,321%	17,67 9%	98,571 %	98,571 %	99,10 7%
E	14,286%	14,28 6%	99,286 %	99,107 %	99,46 4%	10,536%	17,32 1%	99,107 %	98,393 %	99,64 3%	12,857%	20,71 4%	99,464 %	98,750 %	98,92 9%
F	9,464%	15,53 6%	98,571 %	98,214 %	99,46 4%	13,750%	16,78 6%	92,143 %	98,036 %	99,64 3%	10,357%	23,21 4%	98,750 %	98,929 %	94,28 6%
G	13,571%	12,85 7%	96,964 %	98,214 %	97,67 9%	15,714%	10,35 7%	97,500 %	96,964 %	98,39 3%	14,286%	17,50 0%	98,036 %	97,857 %	97,67 9%
H	14,286%	14,28 6%	98,750 %	98,393 %	97,50 0%	9,286%	21,60 7%	98,929 %	98,750 %	98,92 9%	22,679%	17,14 3%	98,929 %	98,750 %	97,14 3%
I	21,250%	11,42 9%	96,607 %	99,464 %	98,57 1%	13,214%	8,036 9%	99,643 %	98,929 %	97,85 7%	18,214%	15,17 9%	98,750 %	98,929 %	98,57 1%
J	14,286%	14,64 3%	98,214 %	97,857 %	98,21 4%	14,107%	15,53 6%	98,214 %	98,393 %	98,92 9%	11,786%	14,82 1%	98,036 %	98,393 %	98,57 1%
K	14,286%	14,28 6%	97,321 %	97,679 %	97,32 1%	15,536%	17,67 9%	95,179 %	98,214 %	95,17 9%	13,750%	27,50 0%	97,679 %	97,500 %	97,14 3%
L	9,107%	13,75 0%	98,214 %	97,321 %	98,39 3%	19,643%	14,10 7%	97,679 %	99,107 %	98,92 9%	18,750%	16,96 4%	99,286 %	93,750 %	98,03 6%
M	15,714%	7,857 %	99,107 %	97,500 %	99,28 6%	20,893%	39,28 6%	98,036 %	99,286 %	99,10 7%	10,000%	77,14 3%	98,750 %	98,214 %	99,28 6%
N	14,286%	14,64 3%	99,464 %	99,286 %	99,82 1%	16,607%	27,50 0%	99,821 %	99,464 %	99,64 3%	14,643%	55,53 6%	99,107 %	99,643 %	99,46 4%
O	14,286%	12,14 3%	99,107 %	99,464 %	99,46 4%	11,964%	20,17 9%	98,571 %	98,750 %	99,10 7%	13,750%	80,00 0%	99,107 %	99,464 %	98,75 0%
P	13,750%	14,28 6%	99,643 %	99,643 %	99,28 6%	13,929%	26,78 6%	98,571 %	99,107 %	99,46 4%	7,321%	47,50 0%	99,286 %	99,464 %	99,46 4%
Q	14,286%	14,28 6%	99,464 %	99,464 %	99,46 4%	6,071%	21,96 4%	98,750 %	99,643 %	99,82 1%	7,679%	50,53 6%	99,464 %	99,821 %	99,82 1%
R	14,286%	16,07 1%	98,929 %	99,286 %	98,92 9%	18,214%	37,32 1%	99,286 %	99,643 %	99,46 4%	14,286%	76,42 9%	100,00 0%	99,643 %	99,10 7%
S	8,929%	14,28 6%	98,571 %	98,393 %	98,03 6%	17,500%	14,46 4%	98,214 %	98,750 %	98,03 6%	17,500%	30,53 6%	98,214 %	98,214 %	97,85 7%
T	14,286%	12,85 7%	96,964 %	97,500 %	97,67 9%	17,679%	20,00 0%	98,393 %	97,321 %	97,32 1%	9,464%	31,07 1%	97,321 %	97,679 %	98,39 3%
U	8,214%	13,75 0%	96,071 %	96,786 %	97,14 3%	15,357%	13,92 9%	98,214 %	98,036 %	98,39 3%	14,821%	17,85 7%	98,036 %	97,679 %	98,21 4%
V	14,286%	12,85 7%	98,214 %	95,000 %	96,25 0%	16,250%	16,25 0%	96,250 %	96,964 %	98,21 4%	15,536%	13,57 1%	97,500 %	98,393 %	96,25 0%
W	14,286%	14,28 6%	96,250 %	96,071 %	96,07 1%	14,821%	15,00 0%	96,964 %	97,321 %	96,60 7%	14,107%	10,89 3%	98,393 %	98,393 %	97,67 9%
X	13,750%	23,75 0%	98,929 %	98,571 %	98,92 9%	14,643%	28,57 1%	98,929 %	98,393 %	98,75 0%	16,071%	42,67 9%	98,393 %	99,464 %	97,85 7%

IV. KESIMPULAN

Dalam hasil uji coba yang telah dilakukan didapatkan hasil bahwa dari 360 percobaan BiLSTM yang telah dikelompokkan ke dalam 24 percobaan (A sampai X) memiliki nilai rata-rata tertinggi pada percobaan R yaitu 71,393% , sedangkan dengan kombinasi fungsi *activation* dan *optimizer* memiliki nilai rata-rata tertinggi pada ReLU dan Adam yaitu 98,579%. Penelitian ini dapat membuktikan bahwa klasifikasi menggunakan suara dapat digunakan terutama pada burung endemik Indonesia yang dilindungi.

DAFTAR PUSTAKA

- [1] H. Alikodra, *Konservasi Sumberdaya Alam dan Lingkungan: Pendekatan Ecosophy Bagi Penyelamatan Bumi*. Gajdiah Mada University Pres, 2012.
- [2] WCS, "World Conservation Strategy Living Resource Conservation for Sustainable Development," 1980.
- [3] KBBI, "Kamus Besar Bahasa Indonesia [Online] - eksploitasi." Accessed: Oct. 20, 2023.
- [4] W. S. Guntur and S. Slamet, "Kajian Kriminologi Perdagangan Ilegal Satwa Liar," *Recidive*, vol. 8, no. 2, 2019, Accessed: May 02, 2024.
- [5] F. F. F. Funay, "Rancang Bangun Aplikasi Pengenalan Suara Burung Dengan Long-Short Term Memory," 2021.

-
- [6] H. Liu, C. Liu, T. Zhao, and Y. Liu, "Bird Song Classification Based on Improved Bi-LSTM-DenseNet Network," in *2021 4th International Conference on Robotics, Control and Automation Engineering, RCAE 2021*, Institute of Electrical and Electronics Engineers Inc., 2021, pp. 152–155.
- [7] Bismiarti, "Maraknya Aktivitas Jual-Beli Satwa Dilindungi di Pasar Burung." Accessed: Nov. 04, 2023.
- [8] Pemerintah Indonesia, "Undang-Undang Nomor 5 Tahun 1990." Accessed: Nov. 04, 2023. [Online]. Available: <https://www.hukumonline.com/pusatdata/detail/371/undangundang-nomor-5-tahun-1990/document>
- [9] A. Rohman *et al.*, "Eksplorasi Satwa Liar di Indonesia (Telaah Ketentuan Pasal 302 KUHP dan UU No. 41 Tahun 2014)," *Laboratorium Syariah dan Hukum*, vol. 4, pp. 81–97, 2023.
- [10] D. Guimarães Da Silva, A. Alvarenga, and M. Meneses, "Comparing LSTM and BLSTM Deep Neural Networks for Power Consumption Prediction: Preliminary studies," May 2023. Accessed: Oct. 25, 2023.
- [11] M. G. Rizky, "Analisis Perbandingan Metode Lstm Dan Bilstm Untuk Klasifikasi Sinyal Jantung Phonocardiogram Tugas Akhir Program Studi S1 Teknik Komputer Oleh," Surabaya, Jun. 2021.
- [12] P. Antoniadis, "Activation Functions: Sigmoid vs Tanh," Baeldung.
- [13] Z. K. Abdul and A. K. Al-Talabani, "Mel Frequency Cepstral Coefficient and its Applications: A Review," 2022, *Institute of Electrical and Electronics Engineers Inc.*