

Original Article

e-ISSN: 2581-0545 - <https://journal.itera.ac.id/index.php/jsat/>

Received 00th January 20xx
 Accepted 00th February 20xx
 Published 00th March 20xx

Open Access

DOI: 10.35472/x0xx0000

Implementasi Model Transfer Learning Arsitektur ConvNeXt untuk Klasifikasi Suara Burung di Taman Nasional Way Kambas

Rizki Adrian Bennovry^{1a}, Alber Analafean², Nabilah
 Andika Fitriani³, Helma Lia Putri⁴, Catherine
 Firdhasari Maulina Sinaga⁵

^a Program Studi Sains Data, Fakultas Sains, Institut Teknologi Sumatera, Sumatera
 Selatan, Lampung 35365

* Corresponding E-mail: rizki.121450073@student.itera.ac.id

Abstract: Deforestation in Sumatra has threatened biodiversity and wildlife habitats, especially birds. The increasing level of forest destruction has resulted in a significant decline in the population and distribution of bird species. This study aims to implement the ConvNeXt architecture transfer learning model in classifying bird sounds in Way Kambas National Park as an effort to monitor and conserve biodiversity. The research method uses a biophony approach, which is a research approach that utilizes natural sound signals from biotic components (living organisms) in a habitat as a data source to build a computational model that is able to recognize environmental patterns and characteristics through spectral analysis of acoustic signals. The acoustic feature extraction process is carried out using Mels Spectrogram, which allows for a visual representation of the frequency characteristics of bird sounds. Processed through the ConvNeXt transfer learning model with two different strategies: augmentation and non-augmentation. Data augmentation techniques are applied to increase the variability and robustness of the model in recognizing various variations of bird sounds. The classification process is carried out by analyzing sound spectrograms and recognizing specific acoustic patterns of each species. The results show that the ConvNeXt model with data augmentation provides significantly better performance than the non-augmentation model. The resulting increase in classification accuracy proves the effectiveness of the transfer learning and data augmentation approaches in identifying bird species based on their vocal characteristics. This study provides an important contribution to the development of biodiversity monitoring technology using artificial intelligence. This method is not only able to identify bird species accurately, but also has the potential to be an effective tool for monitoring ecosystem changes and supporting conservation efforts in Way Kambas National Park.

Keywords: *ConvNeXt, Bird Sound Classification, Transfer Learning, Conservation, Way Kambas*



Abstrak: Deforestasi di Sumatera telah mengancam keanekaragaman hayati dan habitat satwa liar, khususnya burung. Tingkat kerusakan hutan yang terus meningkat mengakibatkan penurunan populasi dan sebaran spesies burung secara signifikan. Penelitian ini bertujuan untuk mengimplementasikan model transfer learning arsitektur ConvNeXt dalam mengklasifikasikan suara burung di Taman Nasional Way Kambas sebagai upaya pemantauan dan konservasi keanekaragaman hayati. Metode penelitian menggunakan pendekatan biofoni, yaitu pendekatan penelitian yang memanfaatkan sinyal suara alamiah dari komponen biotik (organisme hidup) di suatu habitat sebagai sumber data untuk membangun model komputasi yang mampu mengenali pola dan karakteristik lingkungan melalui analisis spektral sinyal akustik. Proses ekstraksi fitur akustik dilakukan menggunakan Mels Spectrogram, yang memungkinkan representasi visual dari karakteristik frekuensi suara burung. Diproses melalui model transfer learning ConvNeXt dengan dua strategi berbeda: augmentasi dan non-augmentasi. Teknik augmentasi data diterapkan untuk meningkatkan variabilitas dan ketangguhan model dalam mengenali berbagai variasi suara burung. Proses klasifikasi dilakukan dengan menganalisis spektrogram suara dan mengenali pola akustik spesifik dari masing-masing spesies. Hasil penelitian menunjukkan bahwa model ConvNeXt dengan augmentasi data memberikan performa yang signifikan lebih baik dibandingkan model non-augmentasi. Peningkatan akurasi klasifikasi yang dihasilkan membuktikan efektivitas pendekatan transfer learning dan augmentasi data dalam mengidentifikasi spesies burung berdasarkan karakteristik suaranya. Penelitian ini memberikan kontribusi penting dalam pengembangan teknologi pemantauan biodiversitas menggunakan kecerdasan buatan. Metode ini tidak hanya mampu mengidentifikasi spesies burung secara akurat, tetapi juga berpotensi menjadi alat yang efektif untuk memantau perubahan ekosistem dan mendukung upaya konservasi di Taman Nasional Way Kambas.

Kata Kunci : ConvNeXt, Klasifikasi Suara Burung, Transfer Learning, Konservasi, Way Kambas.

Pendahuluan

Taman Nasional Way Kambas merupakan salah satu kawasan konservasi paling penting di Indonesia, terletak di Provinsi Lampung, yang memiliki keanekaragaman hayati luar biasa, terutama dalam hal kekayaan populasi burung. Sebagai habitat yang unik dan kompleks, kawasan ini menyimpan berbagai spesies burung yang sebagian di antara nya merupakan satwa endemik dan dilindungi. Namun, upaya identifikasi dan monitoring populasi burung secara tradisional menghadapi sejumlah tantangan signifikan, di antara nya keterbatasan sumber daya manusia, biaya penelitian yang tinggi, serta kompleksitas proses identifikasi yang memerlukan keahlian khusus. (Andyono et al., 2018, 1). Beberapa tahun terakhir, teknologi pengolahan sinyal audio berbasis kecerdasan buatan telah memberikan alternatif baru dalam pengklasifikasian suara burung. Suara burung memberikan informasi unik yang dapat digunakan untuk mengidentifikasi spesies, sehingga memungkinkan pendekatan non-inovatif dalam pemantauan biodiversitas. Akan tetapi, tantangan utama dalam klasifikasi suara burung adalah kompleksitas data audio, seperti variasi frekuensi, amplitudo serta keberadaan noise lingkungan.

Perkembangan teknologi kecerdasan buatan, khususnya dalam ranah deep learning dan machine learning, telah membuka ruang transformasi yang sangat potensial dalam bidang konservasi dan penelitian keanekaragaman hayati. Transfer Learning memungkinkan

pemanfaatan model yang telah dilatih pada dataset besar untuk diaplikasikan pada domain yang lebih spesifik dengan kebutuhan data pelatihan yang lebih sedikit. Salah satu terobosan teknologi yang menjanjikan adalah penggunaan arsitektur jaringan syaraf tiruan modern seperti ConvNeXt, yang telah menunjukkan kemampuan luar biasa dalam mengekstraksi fitur-fitur kompleks dari berbagai jenis data, termasuk data audio. ConvNeXt memiliki keunggulan dalam arsitektur yang mampu menangkap representasi fitur secara mendalam, menjadikannya ideal untuk klasifikasi suara burung yang memerlukan pengenalan pola temporal dan spektral. (Mmileng et al., 2024, 2)

Penelitian R Arif Firmansah et al. tentang Implementasi Transfer Learning pada Pengenalan Citra Rumah Adat Indonesia menggunakan metode CRISP-DM untuk mengklasifikasikan lima kelas rumah adat. Penelitian menguji empat arsitektur model transfer learning: MobileNetV2, VGG16, Xception, dan ResNet50, dengan hasil akurasi bervariasi. Model MobileNetV2 mencapai akurasi tertinggi 0,96%, diikuti ResNet50 (0,94%), Xception (0,92%), dan VGG16 (0,74%). Temuan kunci menunjukkan bahwa performa model machine learning sangat bergantung pada karakteristik dataset spesifik, menegaskan pentingnya eksperimentasi dengan berbagai arsitektur model. Peneliti merekomendasikan pengembangan penelitian selanjutnya dengan memperbesar dataset, menambah variasi kelas, dan mengembangkan proses deployment machine learning yang lebih komprehensif. Penelitian ini signifikan karena mendemonstrasikan efektivitas transfer learning dalam

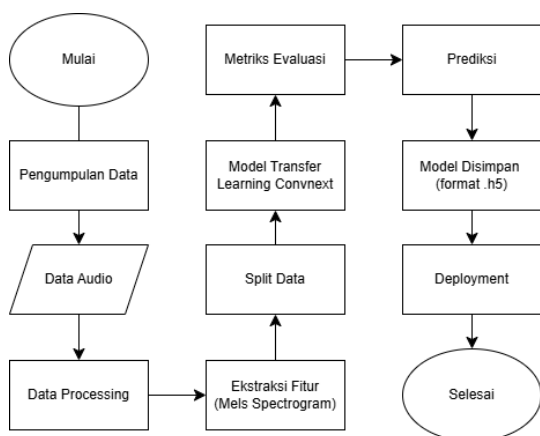
klasifikasi citra warisan budaya Indonesia, membuka peluang aplikasi teknologi machine learning untuk pelestarian budaya. (Firmansah et al., 2023, #)

Penelitian ini fokus pada implementasi model transfer learning berbasis arsitektur ConvNeXt untuk mengklasifikasikan suara burung di Taman Nasional Way Kambas. Pendekatan ini tidak hanya menawarkan solusi teknologi yang inovatif, tetapi juga memberikan prospek signifikan dalam mempercepat dan mengoptimalkan proses identifikasi spesies burung dengan akurasi yang tinggi. Penelitian ini diharapkan dapat memberikan solusi yang efektif dan efisien dalam mendukung upaya konservasi melalui pemantauan keanekaragaman burung, serta mengevaluasi performa model dari aspek akurasi dan efisiensi waktu pelatihan.

Metode

1. Diagram Alir

Penelitian ini dilakukan dengan beberapa tahapan proses yang sistematis untuk mendapatkan data yang relevan dan melakukan analisis secara kuantitatif. Berikut adalah tahapan dan pendekatan yang digunakan:



Gambar 1. Proses Deployment Model ke dalam Sistem Berbasis Website

2. Pengumpulan Data

Data suara burung diperoleh dari database audio burung di seluruh wilayah (<https://xeno-canto.org/>) dengan melakukan *data crawling* menggunakan API Xeno-Canto dan informasi database nama burung diseluruh wilayah Avibase. Data yang dikumpulkan mencakup berbagai kelas suara burung yang ada di Taman Nasional Way Kambas. Penelitian ini difokuskan pada wilayah Taman Nasional Way Kambas untuk mendapatkan daftar burung berdasarkan nama ilmiah, genus, dan spesies. Daftar tersebut digunakan sebagai acuan utama dalam proses klasifikasi. Berdasarkan pada informasi yang terverifikasi dari Avibase, penelitian ini memastikan bahwa data spesies burung yang dianalisis relevan dan valid dalam konteks geografis Way Kambas.

Data dikumpulkan secara Data Crawling dengan memanfaatkan API dari Xeno-Canto, dengan persebaran type audio database xeno-canto. Dataset yang diperoleh yaitu rata-rata 191 audio di seluruh kelas, dengan total keseluruhan data audio sebanyak 1912 audio. Data audio burung pada tabel 2.1.1 dilakukan pemodelan hanya beberapa audio pada masing-masing kelas, dengan menggunakan model transfer learning arsitektur ConvNeXt. Dataset yang digunakan sebagai berikut.

Tabel 1. Dataset Burung Xeno-Canto

Nama Spesies	Nama Ilmiah	Jumlah Data Audio
Kedasi Hitam	<i>Surniculus Lugubris</i>	156
Tekukur Biasa	<i>Spilopelia Chinensis</i>	223
Prenjak Rawa	<i>Prinia Flaviventris</i>	267
Sikatan Kepala Abu	<i>Culicicapa Ceylonensis</i>	273
Perkutut Jawa	<i>Geopelia Striata</i>	133
Punggok Cokelat	<i>Ninox Scutulata</i>	176
Paok Hijau	<i>Pitta Sordida</i>	159
Wiwik Kelabu	<i>Cacomantis Merulinus</i>	201
Cinenen Kelabu	<i>Orthotomus Ruficeps</i>	125
Cekakak Belukar	<i>Halcyon Smymensis</i>	199

3. Preprocessing

Tahap preprocessing dimulai dengan memilih 100 sampel audio secara acak untuk setiap kelas burung. Tujuannya adalah menciptakan dataset yang seimbang dan representatif. Selanjutnya, dilakukan resampling frekuensi untuk menstandarisasi format audio ke frekuensi sampling 16000 Hz. Proses segmentasi audio akan membatasi durasi setiap rekaman menjadi 5 detik menggunakan librosa, yang kemudian dikelompokkan berdasarkan kelas yang selanjutnya disimpan dengan format baru dan setiap segmen diberi nama sesuai kelas suara burung, yang bertujuan menciptakan input model yang konsisten dan terukur, serta mengurangi variabilitas yang tidak perlu dalam data latih. (Simonyan et al., 2014)

4. Ekstraksi Fitur

Fitur data diekstraksi dalam bentuk *Mel Spectrogram* menggunakan transformasi *Fourier* untuk menangkap karakteristik spektral dan temporal dari suara burung. *Mel Spectrogram* adalah metode yang digunakan untuk mengubah sinyal audio mentah menjadi representasi visual berbasis spektrum frekuensi dalam domain waktu. Teknik ini menggunakan skala Mel, yang dirancang menyerupai cara manusia memahami perbedaan frekuensi, dengan memberikan resolusi lebih tinggi pada frekuensi rendah. Dalam proses ini, sinyal audio diproses untuk menghasilkan visualisasi dua dimensi, di mana sumbu waktu, frekuensi, dan intensitas suara ditampilkan dalam bentuk gradasi warna. Representasi ini sangat efektif untuk menangkap karakteristik unik dari pola suara burung, sehingga mempermudah model machine learning dalam mengenali dan mengklasifikasikan suara berdasarkan spesiesnya. Hasil ekstraksi fitur akan menjadi masukan utama untuk model. (Rafliansyah et al., 2024)

Berikut tahapan Pembuatan Mel Spectrogram :

- 1) Transformasi Fourier Cepat (Fast Fourier Transform - FFT) : Mengubah sinyal audio dari domain waktu menjadi domain frekuensi. Rumus dasar FFT :

$$X(k) = \sum_{n=0}^{N-1} x(n) \cdot e^{-j2\pi kn/N}, \quad k = 0, 1, \dots, N-1$$

Dimana $X(k)$ adalah komponen frekuensi, $x(n)$ adalah sinyal input, N adalah panjang sinyal, dan j adalah unit imajiner

- 2) Penerapan Filter Bank Mel :

Frekuensi linear dari hasil FFT diubah menjadi skala Mel menggunakan filter triangular. Konversi dari frekuensi $f(\text{Hz})$ ke frekuensi Mel dilakukan dengan :

$$m = 2595 \cdot \log_{10}\left(1 + \frac{f}{700}\right)$$

filter ini memberikan resolusi lebih tinggi pada frekuensi rendah

- 3) Logaritma Spektral :

Menghitung logaritma dari magnitudo spektrum untuk meniru persepsi manusia terhadap intensitas suara.

- 4) Transformasi Cepat Terbalik (Inverse Discrete Cosine Transform - IDCT) Menyediakan visualisasi dua dimensi (waktu vs frekuensi) dalam bentuk Mel Spectrogram.

Proses ekstraksi fitur dengan Mels Spectrogram dilakukan pada masing-masing kelas. Parameter utama pengaturan Mels-Spectrogram yang digunakan yaitu:

- a. Sample rate (sr) dengan frekuensi : 16.000 Hz
- b. FFT window size (n_{fft}) : 1024 sampel
- c. Hop length : 512 sampel (overlap antar jendela analisis)
- d. Jumlah filter Mel (n_{mels}) : 128

Dengan nilai power hasil Mel-Spectrogram dikonversi ke skala desibel menggunakan *librosa.power_to_db* untuk memudahkan interpretasi. Setelah itu, Mel-Spectrogram divisualisasikan dalam bentuk gambar menggunakan skema warna 'magma'. Setiap file audio disimpan dalam format gambar .png. Tahap ekstraksi fitur mengonversi file audio masing-masing kelas menjadi representatif Mel-Spectrogram.

5. Subsampling

Tahap subsampling akan membatasi jumlah data menjadi 150 sampel dari hasil mel spectrogram menggunakan teknik stratified sampling. Tujuan utama dari proses ini adalah mengoptimasi beban komputasi dan mencegah overfitting, dengan memastikan bahwa setiap kelas tetap terwakili secara proporsional. Metode ini membantu mengendalikan kompleksitas model tanpa mengorbankan keanekaragaman informasi dalam dataset.

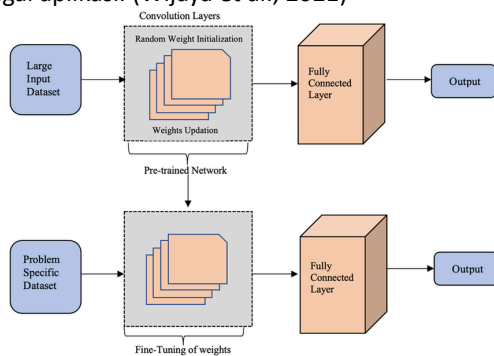
6. Augmentasi Data

Teknik yang digunakan meliputi shear, zoom, flip horizontal, dan rotasi pada gambar mel spectrogram (Simonyan et al., 2014). Augmentasi data dilakukan untuk meningkatkan variabilitas data pelatihan dan ketahanan model terhadap perubahan kecil. Teknik augmentasi mencegah model overfitting dengan variasi tambahan pada data pelatihan (train). Augmentasi data menggunakan *ImageDataGenerator* yang memberikan variasi data melalui

transformasi. Augmentasi hanya diterapkan pada data training, sedangkan data validasi hanya dinormalisasi (*rescale*). Melalui manipulasi ini, model diharapkan dapat lebih robust terhadap variasi pola visual yang merepresentasikan suara burung, mampu mengenali pola akustik burung dalam kondisi yang berbeda, dan mengurangi potensi bias dalam proses klasifikasi. (Toyib et al., 2024, 7).

7. Model Transfer Learning

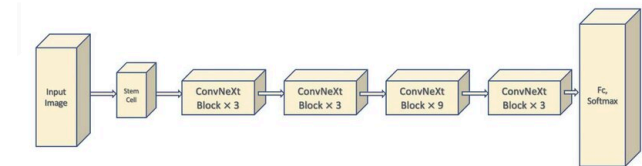
Transfer Learning merupakan pembelajaran mesin dimana sebuah model yang telah dilatih sebelumnya. Transfer learning adalah model yang memanfaatkan pengetahuan yang telah diperoleh dari pelatihan sebelumnya untuk mempercepat proses pembelajaran dan meningkatkan performa model. Transfer learning memiliki keunggulan diantaranya dapat menghemat waktu dan sumber daya dengan melatih model dari awal yang memerlukan banyak waktu pada proses komputasi. Transfer learning juga banyak digunakan dalam machine learning seperti pengenalan objek, pengenalan suara, pemrosesan bahasa alami, dan mendeteksi emosi dalam gambar. Transfer learning ini, teknik yang sangat efektif dalam pengembangan sebuah model machine learning. Terutama, dalam menghadapi keterbatasan data dan sumber daya. Dengan memanfaatkan pengetahuan dari model, dapat mempercepat proses pelatihan dan meningkatkan akurasi pada berbagai aplikasi. (Wijaya et al., 2021)



Gambar 2. Arsitektur Model Transfer Learning

ConvNeXt adalah arsitektur jaringan syaraf konvolusional (CNN) yang dirancang untuk meningkatkan kinerja model dalam menjalankan tugas-tugas pemrosesan gambar, seperti klasifikasi suara burung dan deteksi objek burung. Model ini terinspirasi oleh desain *Vision Transformers* dengan tetap mempertahankan sifat efisien CNN dan memiliki tujuan untuk menggabungkan keunggulan dari kedua pendekatan tersebut. ConvNeXt memiliki kedalaman yang signifikan yang terdiri dari lapisan

konvolusi, lapisan aktivasi, dan lapisan yang terhubung sepenuhnya. ConvNeXt menunjukkan performa yang sangat baik menggunakan *ImageNet*. ConvNeXt juga banyak digunakan dalam berbagai aplikasi pengolahan citra, seperti klasifikasi suara burung, mendeteksi objek burung, maupun segmentasi. (Austin et al., 2024)



Gambar 3. Arsitektur ConvNeXt

Arsitektur ConvNeXt Type Base dipilih sebagai model transfer learning, dengan menggunakan framework *TensorFlow* dan *Keras*. Model akan diinisiasi dengan bobot pra-latih dari *ImageNet*, dan seluruh lapisan akan di-*fine-tuning*. Pendekatan transfer learning memungkinkan pemanfaatan pengetahuan dari model yang sudah dilatih sebelumnya, mengadaptasikannya ke konteks klasifikasi suara burung (Pedregosa & Fabian, 2011, 6). Untuk memastikan performa model yang optimal, optimizer yang digunakan model adalah Adam, yang dapat menyesuaikan learning rate selama proses pelatihan. (Zhang & Zijun, 2018)

8. Performa Model

Model Performa model akan di visualisasikan melalui grafik akurasi dan loss menggunakan Matplotlib dan Seaborn. Grafik akan menampilkan kurva training dan validasi, memberikan wawasan visual tentang proses pembelajaran model, menunjukkan konvergensi, potensi overfitting atau underfitting, serta membantu peneliti dalam menginterpretasi kinerja arsitektur ConvNeXt.

9. Evaluasi Metrik

Evaluasi metrik dilakukan secara komprehensif menggunakan Scikit-learn, mencakup confusion matrix, akurasi, presisi, recall, dan F1-score. Analisis diperluas hingga perhitungan akurasi per kelas untuk memberikan pemahaman yang lebih mendalam terkait performa model dalam mengklasifikasikan setiap spesies burung. Metode ini memungkinkan penilaian kuantitatif yang lebih akurat terhadap kemampuan model. Tambahan rumus-rumus evaluasi metrik seperti presisi (Precision), recall (Sensitivity), dan F1-score digunakan untuk melengkapi analisis performa model (Abadi et al., 2016).

Berikut adalah semua rumus evaluasi metrik yang digunakan dalam Evaluasi Klasifikasi:

- Akurasi

$$Accuracy = \frac{TP+TN}{Total\ Samples}$$

- Presisi

$$Precision = \frac{TP}{TP+FP}$$

- Recall

$$Recall = \frac{TP}{TP+FN}$$

- F-1 Score

$$F1 = 2 \times \frac{Precision \times Recall}{Precision+Recall}$$

Keterangan :

TP = True Positive

FP = False Positive

FN = False Negative

TN = True Negative

10. Prediksi

Penelitian ini menggunakan stratifikasi dengan proporsi 80% untuk data latih (training), 10% untuk validasi, dan 10% untuk pengujian (testing). Teknik pembagian ini memastikan distribusi kelas yang seimbang pada setiap subset data. Pada tahap training, model ConvNeXt akan dilatih menggunakan 80% dataset, yang memungkinkan model untuk mempelajari pola-pola kompleks dalam klasifikasi suara burung. Validasi 10% digunakan untuk melakukan evaluasi berkala selama proses pelatihan, membantu mencegah overfitting dan memastikan generalisasi model. Subset testing 10% akan digunakan untuk mengevaluasi performa akhir model, memberikan penilaian independen dan objektif terhadap kemampuan klasifikasi model pada data yang sama sekali tidak pernah dilihat sebelumnya.

11. Deployment

Tahapan deployment dilakukan dalam implementasi interface model yang dibangun dengan menggunakan framework stramlit. Streamlit merupakan library dari python yang berguna dalam membuat tampilan

user interface dari python (Iwandini et al., 2023). Dimana tahapan deployment memanfaatkan basis pemrograman python, css serta html. Penggunaan streamlit agar bisa model dibangun ditampilkan dalam bentuk user interface yang bisa dimanfaatkan oleh pengguna yang luas.

Hasil dan Pembahasan

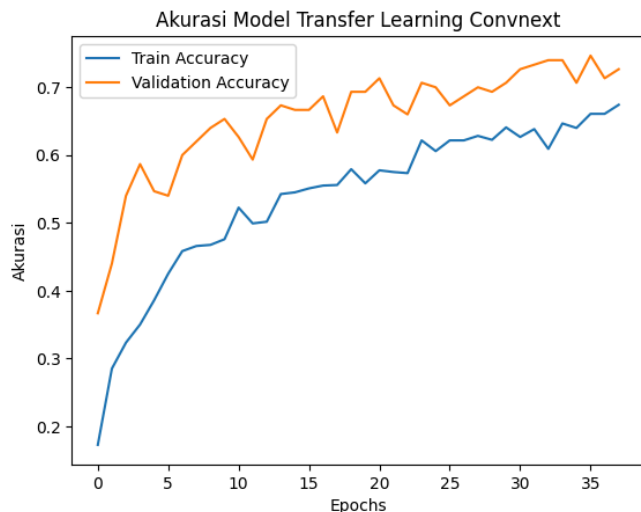
Penelitian ini menggunakan 2 dugaan yaitu model yang menggunakan augmentasi dan model yang menggunakan non augmentasi. Kombinasi parameter yang dilakukan dengan masing masing dugaan tersebut memiliki 2 parameter yang berbeda. Parameter tersebut dilakukan pemodelan dan didapatkan bahwa hasil dengan dugaan tanpa augmentasi pada epoch 25 didapatkan akurasi sebesar 0.70 dan pada epoch 50 didapatkan akurasi sebesar 0.69. Pada dugaan menggunakan augmentasi pada epoch 25 didapatkan akurasi sebesar 0.65 dan pada epoch 50 didapatkan akurasi sebesar 0.71. Berdasarkan parameter yang dijalankan pada model transfer learning arsitektur ConvNeXt didapatkan bahwa model dengan augmentasi pada epoch 50 memiliki akurasi yang cukup tinggi diantara parameter yang lain sehingga didapatkan model augmentasi pada epoch 50 menunjukkan model yang terbaik dibandingkan parameter yang lain. Parameter yang sudah dijalankan dapat dilihat pada Tabel 2.

Tabel 2. Parameter Model Dengan Data Non-Augmentasi & Augmentasi

Batch Size	Learning Rate	Augmentasi Data	Epochs	Akurasi
32	1e-4(0.0001)	Tanpa Augmentasi Data	25	0.70
			50	0.69
	1e-4(0.0001)	Augmentasi Data	25	0.65
			50	0.71

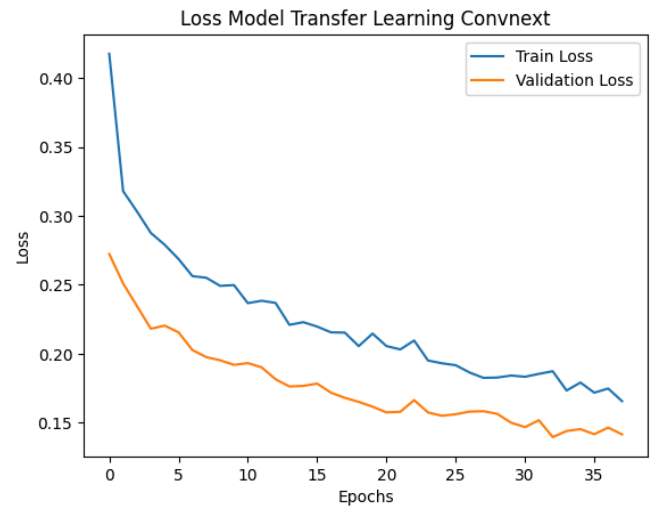
Evaluasi performa model dengan akurasi mencapai 0.71 (71%) yang menunjukkan bahwa model mampu mengklasifikasikan sebagian besar data dengan baik atau performa model yang cukup baik. Augmentasi membantu model untuk *generalize* lebih baik pada data yang belum pernah dilihat sebelumnya, sehingga mengurangi risiko *overfitting*. Performa sebesar 71% sudah cukup baik, dengan model yang dilatih dengan data augmentasi yang memiliki variasi tinggi. Model mampu menangkap pola-pola penting

dalam gambar burung. Dan berikut grafik hasil akurasi pelatihan model transfer learning ConvNeXt pada data augmentasi 50 epoch.



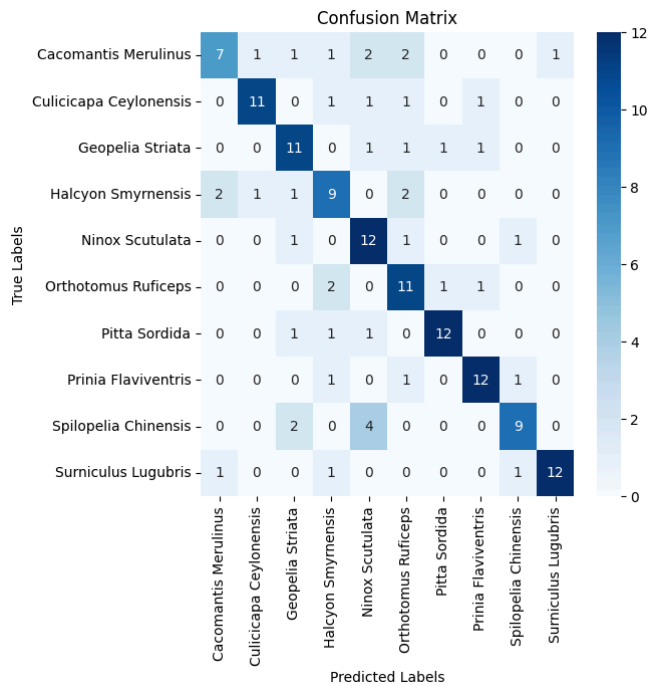
Gambar 4. Akurasi Model Data Augmentasi Epochs 50

Pada gambar 4 Grafik menunjukkan performa model transfer learning ConvNeXt yang dilatih dengan data augmentasi. Pada hasil grafik diatas yang menjadi parameter adalah epoch 50 dengan batch size 32 dan learning rate yang digunakan yaitu 0.0001. Akurasi model pada data pelatihan (train accuracy) semakin meningkat seiring bertambahnya epoch, menandakan model semakin mahir dalam mengklasifikasikan data latih. Akurasi validasi (validation accuracy) juga meningkat secara konsisten, meskipun dengan sedikit fluktuasi. Hal ini mengindikasikan model dapat menggeneralisasi dengan baik. Grafik juga menunjukkan kecenderungan konvergensi antara akurasi train dan validasi, serta stabilitas performa setelah epoch 20. Secara keseluruhan, model ConvNeXt yang dilatih dengan data augmentasi menunjukkan performa yang baik dalam mengklasifikasikan suara burung. Temuan ini memberikan pemahaman tentang kemampuan model dalam konteks klasifikasi suara burung menggunakan transfer learning.



Gambar 5. Loss Model Data Augmentasi dengan Epochs 50

Gambar 5 menunjukkan performa model transfer learning ConvNeXt berdasarkan matrik loss selama pelatihan menggunakan data augmentasi. Nilai loss pada data pelatihan (train loss) menurun secara signifikan seiring bertambahnya epoch, menandakan model semakin baik dalam mempelajari pola data latih. Demikian pula, loss validasi (validation loss) menurun secara konsisten dengan laju penurunan yang hampir seimbang dengan train loss, mencerminkan kemampuan generalisasi model yang baik tanpa indikasi overfitting. Grafik ini juga menunjukkan stabilitas performa model setelah epoch ke-20. Secara keseluruhan, model ConvNeXt yang dilatih dengan data augmentasi menunjukkan performa yang memadai untuk mengklasifikasikan suara burung, dengan hasil yang stabil dan konvergen di akhir pelatihan. Temuan ini memberikan gambaran yang mendalam tentang efisiensi metode transfer learning dalam klasifikasi suara burung di Taman Nasional Way Kambas.



Gambar 6. Evaluasi Model

Gambar 6 menampilkan evaluasi model menggunakan *confusion matrix* yang mengevaluasi kinerja model ConvNeXt dalam mengklasifikasikan suara burung di Taman Nasional Way Kambas. Model menunjukkan akurasi tinggi pada kelas seperti *Ninox Scutulata* dan *Orthotomus Ruficeps*, namun terdapat beberapa kesalahan prediksi, seperti *Halcyon Smyrnensis* yang terkadang diklasifikasikan sebagai kelas lain. Kesalahan ini disebabkan oleh kemiripan pola suara antar spesies atau distribusi data yang tidak merata. Matriks ini membantu mengidentifikasi area perbaikan untuk meningkatkan akurasi model.

Dari *confusion matrix* tersebut, diperoleh *classification report* yang terlihat pada Tabel 3. dibawah yang menunjukkan model transfer learning pada data augmentasi dengan epochs 50 memiliki akurasi cukup tinggi sebesar 0.71 serta nilai *precision*, *recall*, dan *f1-score* mendekati 1 yang menunjukkan bahwa model bekerja dengan cukup baik.

Tabel 3. Classification Report

Class	Precision	Recall	f1-score	support
0	0.70	0.47	0.56	15
1	0.85	0.73	0.79	15

2	0.65	0.73	0.69	15
3	0.56	0.60	0.58	15
4	0.57	0.80	0.67	15
5	0.58	0.73	0.65	15
6	0.86	0.80	0.83	15
7	0.80	0.80	0.80	15
8	0.75	0.60	0.67	15
9	0.92	0.80	0.86	15
accuracy		0.71		150
macro avg	0.72	0.71	0.71	150
weighted avg	0.72	0.71	0.71	150

Pada tabel 4 menunjukkan hasil akurasi model pada masing-masing kelas yang menunjukkan performa yang bervariasi antar kelas. Kelas dengan akurasi tertinggi (*Ninox Scutulata*, *Pitta Sordida*, *Prinia Flaviventris*, dan *Surniculus Lugubris*) memiliki akurasi sebesar 80% yang menunjukkan kemampuan model dalam mengenali fitur-fitur penting dari spesies burung. Akan tetapi, terdapat kelas dengan akurasi rendah seperti *Cacomantis Merulinus* sebesar 46.67% yang menunjukkan bahwa fitur pada kelas ini lebih sulit dipelajari oleh model. Secara keseluruhan, hasil akurasi masing-masing kelas menunjukkan bahwa model memiliki performa yang cukup baik.

Tabel 4. Hasil Akurasi Per Class

Name Class	Total Akurasi
<i>Cacomantis Merulinus</i>	46.67%
<i>Culicicapa Ceylonensis</i>	73.33%
<i>Geopelia Striata</i>	73.33%
<i>Halcyon Smyrnensis</i>	60%

<i>Ninox Scutulata</i>	80%
<i>Orthotomus Ruficeps</i>	73.33%
<i>Pitta Sordida</i>	80%
<i>Prinia Flaviventris</i>	80%
<i>Spilopelia Chinensis</i>	60%
<i>Surniculus Lugubris</i>	80%

Setelah, dilakukan evaluasi matriks, selanjutnya dilakukan tahapan menguji model (testing) menggunakan data yang baru. Diperoleh hasil prediksi dari model yang dibangun dengan menggunakan data baru, yang bisa terlihat pada tabel 4. bahwa tebakan kelasnya sudah benar walaupun akurasi model yang dibangun pada tabel 5. telah menunjukkan hasil yang cukup baik.

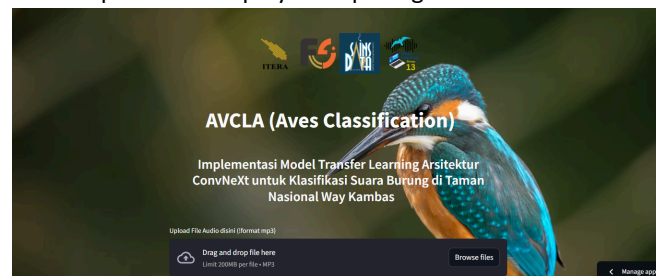
Tabel 5 Hasil Prediksi

Gambar	Actual	Prediksi	Validasi Klasifikasi
	<i>Surniculus Lugubris</i>	<i>Surniculus Lugubris</i>	Benar
	<i>Spilopelia Chinensis</i>	<i>Spilopelia Chinensis</i>	Benar
	<i>Prinia flaviventris</i>	<i>Prinia flaviventris</i>	Benar
	<i>Culicicapa Ceylonensis</i>	<i>Culicicapa Ceylonensis</i>	Benar

*Geopelia striata**Geopelia striata*

Benar

Setelah diperoleh model terbaik dari beberapa parameter, dilakukan tahapan deployment menggunakan streamlit, agar model bisa melakukan prediksi secara tampilan berbasis website. Tahapan deployment dibangun dengan basis pemrograman python, html serta css. Dan diperoleh tampilan hasil deployment pada gambar.



Kesimpulan

Implementasi dan pengujian model transfer learning arsitektur ConvNext menggunakan audio suara burung di Taman Nasional Way Kambas di Provinsi Lampung diperoleh hasil signifikan dapat mengenali dan klasifikasi objek burung. Implementasi proses dengan data augmentasi 50 epochs memiliki peran yang penting dalam model memperoleh performa model dan akurasi yang cukup baik. Dari model dibangun dengan menggunakan beberapa pengaturan parameter, diperoleh bahwa model memiliki akurasi yang cukup baik. Ketika pada pengaturan parameter model yaitu dengan batch size sebesar 32, learning rate 0.0001 dan epochs sebesar 50, diperoleh akurasi model sebesar 0.71 yang menunjukkan model telah cukup baik. Evaluasi model yang cukup tinggi, bahwa model bisa diimplementasikan secara aplikatif seperti dengan mengembangkan aplikasi sistem deteksi klasifikasi suara burung pada kawasan area konservasi seperti cagar alam, suaka margasatwa, taman nasional bahkan geopark.

References

Andyono, Gebyar, Marsono, Djoko, Sadono, Ronggo, Imron, & M Ali. (2018). andyono2018analysis. *The analysis on the stakeholders of conflict mitigation*

in the Way Kambas National Park, Lampung, 22, 15-28.

- Firmansah, R Arif, Santoso, Handri, Anwar, & Agus. (2023). firmansah2023transfer. *Transfer Learning Implementation on Image Recognition of Indonesian Traditional Houses, 4*, 1469--1478.
- Mmileng, Outlwile, Whata, Albert, Olusanya, Micheal, Mhlongo, & Siyabonga. (2024). mmileng2024application. *Application of ConvNeXt with Transfer Learning and Data Augmentation for Malaria Parasite Detection in Resource-Limited Settings Using Microscopic Images, 2024-10*.
- Simonyan, K., & Zisserman, A. (2014). *Very Deep Convolutional Networks for Large-Scale Image Recognition*. arXiv preprint arXiv:1409.1556.
- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., ... & Zheng, X. (2016). *TensorFlow: A system for large-scale machine learning*. In *12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16)* (pp. 265-283).
- Pedregosa, F., et al. (2011). *Scikit-learn: Machine Learning in Python*. *Journal of Machine Learning Research, 12*, 2825-2830.
- Abadi, M., et al. (2016). *TensorFlow: A System for Large-Scale Machine Learning*.