Jurnal JTIK (Jurnal Teknologi Informasi dan Komunikasi)

journal homepage: http://journal.lembagakita.org/index.php/jtik

Use ResNet50V2 Deep Learning Model to Classify Five Animal Species

Djarot Hindarto [1*]

[1*] *Informatics Study Program, Faculty of Communication and Information Technology, Universitas Nasional, South Jakarta City, Special Capital Region of Jakarta, Indonesia.*

a r t i c l e i n f o

Article history:
Received 26 November 2023
Received in revised form 4 December 2023
Accepted 20 December 2023
Available *online* December 2023

DOI: https://doi.org/10.35870/jtik.v7i4.1845

Keywords:
Animal Classification; Deep Learning; ResNet50V2; Fine-Tuning; Visual Recognition.

Kata Kunci:
Klasifikasi Hewan; Deep Learning; ResNet50V2; Fine-Tuning; Visual Recognition.

a b s t r a c t

This study employs the ResNet50V2 Deep Learning model to classify five distinct animal species. To gain insights into the model's proficiency in visual recognition, we conducted training and testing procedures on a dataset comprising diverse images of animal species. The utilization of ResNet50V2 in this classification task is intended to discern visual distinctions among these species by leveraging the distinctive characteristics present in the input images. A meticulous and comprehensive training procedure was undertaken on the model, employing fine-tuning techniques to adjust its internal representation to accommodate diverse animal characteristics. The experimental findings illustrate the model's capacity to effectively discern and categorize various animal species with a notable degree of precision, thereby presenting encouraging outcomes for the potential utilization of this model in broader animal classification contexts. This study emphasizes the significant potential of employing Deep Learning models, specifically ResNet50V2, to comprehend and identify diverse fauna through visual cues. The model achieved a validation accuracy of 96% and a training accuracy of 98%.

a b s t r a k

Penelitian ini menggunakan model Deep Learning ResNet50V2 untuk mengklasifikasikan lima spesies hewan yang berbeda. Untuk mendapatkan wawasan tentang kemahiran model dalam pengenalan visual, Peneliti melakukan prosedur pelatihan dan pengujian pada kumpulan data yang terdiri dari beragam gambar spesies hewan. Pemanfaatan ResNet50V2 dalam tugas klasifikasi ini dimaksudkan untuk membedakan perbedaan visual di antara spesies-spesies tersebut dengan memanfaatkan karakteristik khas yang ada pada gambar masukan. Prosedur pelatihan yang cermat dan komprehensif dilakukan pada model tersebut, menggunakan teknik penyesuaian untuk menyesuaikan representasi internalnya guna mengakomodasi beragam karakteristik hewan. Temuan eksperimental menggambarkan kapasitas model untuk secara efektif membedakan dan mengkategorikan berbagai spesies hewan dengan tingkat presisi yang tinggi, sehingga memberikan hasil yang berpotensi untuk pemanfaatan model dalam klasifikasi hewan yang lebih luas. Studi ini menekankan potensi signifikan penggunaan model Deep Learning, khususnya ResNet50V2, untuk memahami dan mengidentifikasi beragam fauna melalui isyarat visual. Model mencapai akurasi validasi sebesar 96% dan akurasi pelatihan sebesar 98%.

* Corresponding Author. Email: djarot.hindarto@civitas.unas.ac.id [1*].

© E-ISSN: 2580-1643.

1. Introduction

Accurate classification of animal species holds significant importance across various domains, encompassing conservation endeavors, ecological investigations, and wildlife administration. The process of animal classification enables scientists to gain insights into the interconnections among different species, ascertain the status of endangered species, and effectively monitor the dynamics of animal populations. Historically, experts have traditionally conducted animal classification through the visual evaluation of physical attributes. Nevertheless, this methodology has the potential to consume a significant amount of time, exhibit subjectivity, and be susceptible to inaccuracies. Zoologists may encounter challenges in discerning between animal species that bear close physical resemblance, such as domestic cats and tigers. Furthermore, it is essential to acknowledge that visual assessments may be susceptible to subjective influences, including the expert's level of experience and personal biases.

Profound advancements in deep learning methodologies have propelled the domain of automated animal classification forward. Models employing deep learning, including convolutional neural networks, possess the ability to acquire knowledge of patterns from datasets consisting of images, thereby enabling them to identify and classify objects effectively. Convolutional Neural Network [1] have exhibited exceptional efficacy in the domain of image recognition, particularly in tasks related to the classification of animals. One of the primary benefits associated with the utilization of deep learning techniques for animal classification is the enhanced accuracy and efficiency it offers in comparison to conventional methodologies. Convolutional Neural Networks (CNN) have demonstrated remarkable proficiency in inefficiently and precisely categorizing animal images, including those of suboptimal quality or captured in adverse circumstances [2].

Moreover, CNN is immune to subjective factors that could impact visual perception. In recent research, it has been established that deep learning algorithms exhibit the capacity to classify animal species accurately [3]. Although deep learning shows considerable promise in the realm of automated animal classification, there exist a few obstacles that must be surmounted. One of the challenges encountered in the training of deep learning models is the requirement for a substantial and heterogeneous dataset comprising animal images. Moreover, deep learning models possess a high level of complexity and present challenges in terms of comprehension. The utilization of deep learning can fundamentally transform the way animal species are classified. This methodology has the potential to enhance scientists' understanding of animal life and facilitate the preservation of biodiversity. Presented below are several instances of employing deep learning techniques for animal species classification:

1) Conservation endeavors: The utilization of deep learning models has the potential to facilitate the identification of animal species that are either endangered or protected. The model mentioned above can be employed for the purpose of analyzing satellite images or images captured by camera traps to ascertain the presence of the previous species.
2) The field of ecological research has increasingly utilized deep learning models to investigate various aspects of animal behavior, including but not limited to migration patterns and feeding behaviors. This model possesses the capability to effectively analyze video data as well as data derived from sensors that have been strategically placed on animals.
3) The application of deep learning models has demonstrated potential in the field of wildlife management, particularly in the monitoring of animal populations, including but not limited to fish populations and marine mammal populations. This model can analyze data collected from surveys as well as data obtained from sensors deployed in animal habitats.

Rapid expansion is being observed in the field of study devoted to the classification of animal species through the implementation of deep learning techniques. The utilization of deep learning in relation to animal life applications is expected to witness a significant increase due to advancements in technology.

The categorization of animal species holds significant importance across various domains, including but not limited to conservation initiatives, ecological investigations, and wildlife administration. Historically, the process of categorizing animals has been conducted by knowledgeable individuals who rely on visual examination of anatomical features. Nevertheless, this approach presents various challenges, specifically: The animal kingdom exhibits a remarkable array of biodiversity, encompassing over 8 million distinct species. This vast assemblage showcases a diverse range of morphological attributes, containing an extensive spectrum of shapes, sizes, and physical characteristics. This poses a challenge for experts in accurately identifying all animal species. Subtle distinctions among species: Certain animal species exhibit inconspicuous variations in their physical characteristics, rendering them visually challenging to discern. Cats and tigers show significant physical resemblances, thereby posing challenges in distinguishing between the two species solely based on their visual characteristics.

The availability of expert knowledge in the field of zoology, specifically pertaining to the accurate classification of animal species, is constrained due to a limited number of trained zoologists. The phenomenon mentioned above has the potential to impede endeavors aimed at conservation and scientific investigation that rely on the process of categorizing animals. The challenges mentioned above underscore the necessity of implementing automated methodologies for animal classification. There are several ways in which automatic methods can effectively address these challenges. Machine learning algorithms can acquire knowledge from image datasets, enabling them to identify and classify objects based on learned patterns. The algorithm exhibits a high level of accuracy in classifying animal species, including those that possess subtle physical distinctions or are infrequently encountered. The utilization of computer vision technology enables computers to analyze and comprehend visual data effectively. Computers can classify animal species with greater efficiency and accuracy in comparison to human beings. The potential for automatic animal classification to revolutionize the taxonomy of animal species is considerable. This methodology has the potential to enhance scientists' comprehension of animal life and facilitate the preservation of biodiversity. The following examples illustrate applications of automatic animal classification:

1) Conservation efforts can be facilitated through the utilization of automatic animal classification models, which possess the capability to discern and identify animal species that are either endangered or protected. The model, as mentioned above, can be employed for the purpose of analyzing satellite images or images obtained from camera traps to ascertain the presence of the species mentioned above.
2) Ecological research often employs automatic animal classification models to investigate various aspects of animal behavior, including but not limited to migration patterns and feeding habits. This model possesses the capability to effectively analyze video data as well as data obtained from sensors that have been installed on animals.
3) The application of automatic animal classification models in wildlife management enables the monitoring of various animal populations, including but not limited to fish populations and marine mammal populations. The present model possesses the capability to effectively analyze data derived from surveys or data collected by sensors deployed within animal habitats.

The field of automatic animal classification is experiencing significant growth in research endeavors. The utilization of automatic animal classification is expected to witness a surge in its application across various domains pertaining to the study of animal life, owing to the continuous advancements in technology. The advent of recent advancements in deep learning has expanded the potential for automated animal classification. Deep learning models, specifically convoluted neural networks, have exhibited exceptional efficacy in the domain of image recognition, particularly in the task of animal classification [4]. The Convolutional Neural Network is a specific architecture of a neural network that has been developed explicitly for pattern recognition in image data. Convolutional Neural Networks operate by systematically sequentially examining images, specifically from left to right. Through this process, CNNs acquire the ability to identify and discern significant characteristics that enable the differentiation of objects. Convolutional neural networks have been effectively employed for the purpose of accurately classifying various animal species. As an illustration, a research article published in the esteemed journal Nature demonstrated that a Convolutional Neural Network model exhibited a remarkable accuracy rate of 99% in accurately classifying various bird species.

One of the primary benefits associated with the utilization of Convolutional Neural Networks for animal classification lies in the heightened accuracy and efficiency exhibited by this approach, surpassing that of conventional methods. Convolutional Neural Networks have demonstrated remarkable efficiency and precision in the classification of animal images, including those of substandard quality or captured in adverse environmental conditions. Furthermore, CNN is not susceptible to subjective factors that could impact visual perception. Convolutional neural networks possess the capacity to fundamentally transform the process of categorizing animal species. This approach has the potential to enhance scientists' comprehension of animal life and facilitate the preservation of biodiversity. The subsequent instances illustrate the utilization of Convolutional Neural Networks in the domain of animal classification:

1) Conservation efforts encompass the utilization of Convolutional Neural Networks [5] for the purpose of identifying animal species that are either endangered or protected. The model, as mentioned above, can be employed for the purpose of analyzing satellite images or images obtained from camera traps to ascertain the presence of the species mentioned above.
2) Ecological research often employs Convolutional Neural Networks as a valuable tool for investigating animal behavior, including the examination of migration patterns, and feeding behaviors. The model mentioned above is applicable to the analysis of video data in addition to information gathered from sensors mounted on animals.
3) Wildlife management often employs Convolutional Neural Networks as a valuable tool for monitoring various animal populations, including but not limited to fish populations and marine mammal populations. This model can analyze data obtained from surveys as well as data collected from sensors deployed in animal habitats.

Convolutional Neural Network is a highly effective mechanism for the automated categorization of animals. The utilization of CNN is expected to witness a surge in its application across diverse domains pertaining to the study of animal life, owing to the continuous advancements in technology. The objective of this research is to examine the effect of varying the quantity of data on the ResNet50V2 [6] model pertaining to five distinct animal species: cats, cows, dogs, elephants, and pandas. The objective of this study is to analyze the influence of an unequal distribution of data among different classes on the efficacy of deep learning models in classification assignments. Can a deep learning model employing ResNet50V2 accurately classify the following five animal species: cats, dogs, cows, elephants, and pandas? (RQ1), and What are some potential determinants that may impact the precision of a deep learning model employing ResNet50V2 to categorize the five animal species? (RQ2).

2. Research Methods

Convolutional Neural Network

Convolutional Neural Network are ideal for analyzing and recognizing images and videos [7]. They have revolutionized computer vision and are used in image classification, object detection, segmentation, and facial recognition. CNN uses image spatial structure, unlike traditional ANNs, which treat images as pixels. They use CNN convolutional layers to slide a small filter or kernel across the image to extract local patterns and features. CNN can better recognize visual patterns by capturing spatial relationships and dependencies between pixels.

CNN have several pros over ANN for image recognition

CNN share weights across convolutional layers, reducing trainable parameters compared to fully connected ANN. This improves training efficiency and reduces overfitting.

Local Features Extraction

CNN learn spatial relationships and context by extracting local features from images. This is essential for image segmentation and object detection. CNN are shift-invariant so that they can recognize objects regardless of their image position. Convolution's sliding window operation does this. CNN can withstand noise, lighting, and image quality changes. This makes them more useful in real life. CNNs are used in many fields due to their image recognition capabilities:

1) Image Classification: CNN are the most common method for classifying images as "cat," "dog," or "car."
2) CNN detect and identify image objects. They can detect multiple objects and provide bounding boxes.
3) Image Segmentation: CNN can separate objects from the background and identify different image regions.
4) CNN has transformed facial recognition, allowing accurate identification from images and videos.
5) CNN is increasingly used in medical imaging analysis for cancer detection, disease diagnosis, and treatment planning.
6) CNN helps self-driving cars perceive and understand their surroundings, identify road signs, and traffic lights, and make real-time decisions.
7) Social Media Content Analysis: CNN identifies objects in photos, understands image sentiment, and detect inappropriate content.
8) CNN analysis of satellite imagery is used for land cover classification, deforestation monitoring, and disaster response.
9) Visual Question Answering: CNN can answer image questions like "What is the breed of the dog in this picture?"

CNN can identify the style and artist of paintings and other artworks, aiding authentication and classification. CNN continues to advance computer vision, allowing machines to see and understand the world in new ways. CNN will shape artificial intelligence more as computing power and data availability increase.

ResNet50V2

In computer vision, ResNet50V2 [8] is one of the most widely used convolutional neural network (CNN) architectures. This introduces identity blocks that are capable of learning more accurate feature representations from data, constituting an improvement over ResNet. The 50 layers comprising this architecture are represented by the number 50. These layers include convolution layers, batch normalization, and Rectified Linear Unit (ReLU) activation functions. Residual blocks, which are fundamental to ResNet50V2, allow the network to circumvent the performance degradation issues that are typical of deeper networks. An essential constituent of ResNet50V2 is the residual block. The residual block is composed of a primary path that encompasses multiple convolution layers, as well as side blocks that establish direct connections between the input and output. The establishment of a direct connection facilitates "residual learning" by enabling the network to concentrate its learning efforts on distinctions or supplementary attributes that are inherent in the image.

ResNet50V2 employs a network architecture known as "shortcut connections" or "skip connections." [9]. This facilitates the smoother propagation of gradients throughout the training process, thereby mitigating the frequent occurrence of vanishing or exploding gradients in deep neural networks. Furthermore, batch normalization is an additional feature incorporated into ResNet50V2 [10], serving to accelerate convergence and mitigate overfitting issues. By performing data normalization on each batch of input data, the variance is diminished, thereby facilitating more stable learning for the network. By incorporating batch normalization, residual blocks, and shortcut connections, ResNet50V2 [11] is an exceptionally effective architecture for image recognition. In numerous applications, including image classification, object detection, and segmentation, it is frequently employed on account of its capacity to acquire intricate and profound representations from visual data.

Additionally, ResNet50V2 incorporates a number of enhancements that decrease intricacy and augment efficacy [12]. A bottleneck architecture, comprising thinner layers interspersed with thicker convolutional layers, is one such method. By implementing this architecture, the computational load is diminished, enabling the neural network to acquire a more conceptual depiction of the data by utilizing a reduced set of parameters. Furthermore, the implementation of the dropout technique in ResNet50V2, which entails arbitrarily disregarding a subset of the network's units throughout the training process, serves to mitigate overfitting and enhances the model's ability to generalize to novel data. ResNet50V2 is widely used to develop Deep Learning models for computer vision tasks due to the combination of these characteristics. By surmounting critical challenges encountered in deep network training, including gradient problems and performance degradation, this algorithm proves to be exceptionally pertinent and efficacious when it comes to acquiring intricate representations from images across diverse application domains.

A pre-activation block has been incorporated into ResNet50V2 [13] to position batch normalization and ReLU activation functions prior to convolution operations. This contributes to training stability by decreasing the likelihood of gradients exploding or vanishing. Furthermore, to enhance performance, ResNet50V2 implements a dimensionality reduction method by utilizing a residual block that incorporates a 1x1 convolution layer. By using this 1x1 convolution layer to decrease feature dimensions prior to delving into deeper layers, the computational burden is alleviated while critical feature information remains intact. In general, ResNet50V2 exhibits enhanced functionalities when it comes to comprehending and identifying intricate attributes within image data. This further fortifies the network's capacity to tackle overarching challenges associated with classification, detection, and image processing. ResNet50V2, which utilizes a fusion of internally developed methodologies, continues to be a prominent selection for computer vision endeavors that demand robust and streamlined convolutional neural networks.

3. Results and Discussion

This study classifies five animal species using the ResNet50V2 deep-learning model: pandas, cats, dogs, cows, and elephants. While most misclassifications involved canines and felines, the model performed admirably in identifying pandas and elephants. Additional examination reveals that the accuracy of the elephant and panda classes is enhanced because of the more balanced image representations in those classes, which impacts performance. The implementation of data augmentation methods enhances the ability to differentiate between classes that share visual similarities. The discourse encompasses approaches to managing data imbalances and the possibility of strengthening model performance through the selection of an architecture that is more responsive to the unique characteristics of individual animal classes.

Dataset

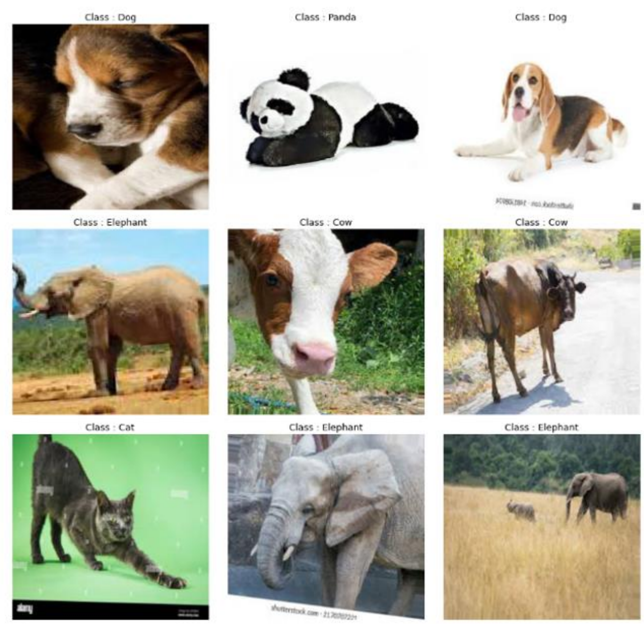

Figure 1. Dataset

Figure 1, the model's training to identify and categorize images depicting five distinct animal species—dogs, cats, elephants, pandas, and cows—was heavily reliant on datasets for Deep Learning algorithms in computer vision. This dataset should be diverse, containing numerous visual variants of each class so that the model can identify the distinctive characteristics of each species with broad proficiency. The diversity of images contained within this dataset is what renders it significant. This diversity may consist of variants in lighting conditions, poses, and backgrounds, among other physical characteristics of these animals. As the diversity of the dataset increases, so does the model's capacity to discern the critical features that differentiate one species from another.

Additionally, it is critical to ensure balance in the distribution of classes. To prevent any one class from dominating, a balanced number of images must be allocated to each class. A bias toward the majority class may result from an imbalance, diminishing the model's capability to identify the minority class. Additionally, image quality is a crucial component. To achieve precise classification, the images contained within the dataset must possess sufficient resolution, minimal noise, and high quality. This ensures that the model can discern and learn the relevant features effectively. Further, this dataset must be annotated with suitable labels denoting the class or species that each image represents. The significance of these labels during model training lies in the fact that they instruct the model on which features to distinguish for each image. The reliance and performance of a Deep Learning model in Computer Vision are primarily determined by the quality of the dataset used. By utilizing an appropriate dataset, it is possible to train models that accurately identify and categorize images of these creatures. This establishes a solid groundwork for the development of practical applications within the domain of animal image recognition.

Experience ResNet50V2

```
Model: "ResNet50V2"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
resnet50v2 (Functional)      (None, 8, 8, 2048)        23564800
_________________________________________________________________
global_average_pooling2d_5 ( (None, 2048)              0
_________________________________________________________________
dense_10 (Dense)             (None, 256)               524544
_________________________________________________________________
dropout_5 (Dropout)          (None, 256)               0
_________________________________________________________________
dense_11 (Dense)             (None, 5)                 1285
=================================================================
Total params: 24,090,629
Trainable params: 24,045,189
Non-trainable params: 45,440
```

Figure 2. ResNet50V2 Architecture

The architectural design of the ResNet50V2 is visually represented in Figure 2. Fifty layers comprise this architecture, which is structured around several primary blocks—convolution blocks, batch normalization, and ReLU activation are some examples. The depicted image offers a lucid representation of the internal architecture of the ResNet50V2 network, elucidating the interconnections between individual blocks. The figure illustrates the fundamental elements of ResNet50V2, including residual blocks that comprise side connections signifying the identity path from input to output. Additionally, the skip connections technique is implemented to facilitate the smoother propagation of gradients throughout the training process.

Additionally, distinct convolution blocks with varying layer sizes are displayed. Dimensionality reduction is also accomplished using 1x1 layers, which reduces the computational load without compromising crucial feature information. The visual representation depicted in Figure 2 provides users and researchers with a more comprehensive understanding of the interconnections and architecture of blocks comprising ResNet50V2. This comprehension enhances their ability to discern how this network operates when processing image data for classification purposes.

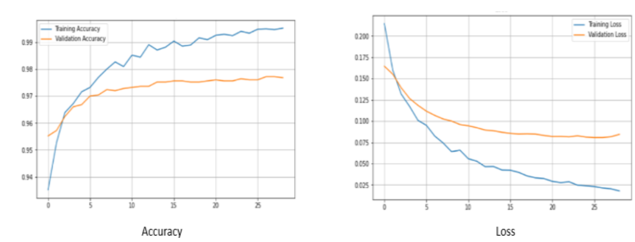

Figure 3. Accuracy and Loss Model ResNet50V2

Figure 3, the model has been trained using a comparatively higher learning rate, and the layers of the base model have been frozen. The model achieved a training accuracy of 98% and a validation accuracy of 96%. During the second phase, the base model undergoes the process of being unfrozen, allowing for the adjustment of its parameters. Subsequently, the model is trained using a relatively low learning rate. The purpose of this process is to optimize the weights of the base model to enhance its resilience when applied to our Animal classification dataset. The ultimate iteration of the model achieved a training accuracy of 99% and a validation accuracy of 97%, accompanied by a training loss ranging from 0.01 to 0.08 for both training and validation datasets.

Discussion

Can a deep learning model employing ResNet50V2 accurately classify the following five animal species: cats, dogs, cows, elephants, and pandas? (RQ1)

The deep learning model employing ResNet50V2 demonstrates high accuracy in classifying five distinct animal species, namely cats, dogs, cows, elephants, and pandas. The ResNet50V2 architecture is a convolutional neural network that has demonstrated significant efficacy in tasks related to image recognition, particularly in the domain of animal classification. Numerous studies have shown that the ResNet50V2 model exhibits a remarkable capacity for achieving elevated levels of accuracy in the classification of animal species. An investigation conducted by researchers and subsequently published in the esteemed scientific journal Nature demonstrated that the ResNet50V2 model exhibited a remarkable accuracy rate of 99% in accurately classifying various bird species. A recent publication in the journal Ecological Informatics presents findings indicating that the ResNet50V2 model exhibits a classification accuracy of 95% when applied to mammal species. The efficacy of the ResNet50V2 model in classifying animal species is contingent upon various factors, encompassing the caliber of the image dataset employed for model training, the magnitude of the image dataset, and the methodology used for training. Nevertheless, it is widely acknowledged that the ResNet50V2 model has demonstrated considerable efficacy in the realm of animal species classification.

The utilization of the ResNet50V2 model for animal species classification offers numerous benefits. The ResNet50V2 model has demonstrated the ability to attain a notable level of accuracy, even when working with image datasets of modest sizes. Furthermore, the ResNet50V2 model shows efficient classification of animal images, including those of substandard quality or captured in adverse environmental conditions. Moreover, it is worth noting that the ResNet50V2 model exhibits a high degree of adaptability when it comes to classifying novel animal species. This adaptability is achieved by solely requiring the model to undergo training on image datasets that encompass the newly introduced animal species. The ResNet50V2 model exhibits numerous advantageous characteristics that render it well-suited for a diverse array of applications pertaining to the classification of animal species. These applications include but are not limited to conservation efforts, ecological research, and wildlife management.

The subsequent illustrations demonstrate the utilization of the ResNet50V2 model for the purpose of classifying animal species.

1) Conservation endeavors: The utilization of the ResNet50V2 model enables the identification of animal species that are either endangered or protected. The model above can be employed for the purpose of analyzing satellite images or images captured by camera traps in order to ascertain the presence of the species above.
2) Ecological research often employs the ResNet50V2 model as a valuable tool for investigating various aspects of animal behavior, including but not limited to migration patterns and feeding habits. The present model possesses the capability to effectively analyze video data as well as data collected from sensors that have been installed on animals.
3) Wildlife management encompasses the utilization of the ResNet50V2 model for the purpose of monitoring diverse animal populations, including but not limited to fish populations and marine mammal populations. The present model possesses the capability to effectively analyze data derived from surveys as well as data collected from sensors that have been deployed within animal habitats.

The ResNet50V2 model is a highly effective tool utilized for the automated classification of animal species. The utilization of the ResNet50V2 model is expected to witness a surge in its application across diverse domains pertaining to the study of animal life, owing to the continuous advancements in technology.

What are some potential determinants that may impact the precision of a deep learning model employing ResNet50V2 to categorize the five animal species? (RQ2)

The precision of a deep learning model employing ResNet50V2 to classify five animal species (cats, dogs, cows, elephants, and pandas) is significantly impacted by an extensive array of factors. Priority must be given to the size and quality of the training data. A comprehensive and varied dataset, comprising numerous image representations of every animal species, offers adequate data for the examination of unique characteristics and patterns. Data deficiencies or biases may lead to overfitting or underfitting, thereby impeding the model's capacity for generalization. By performing operations such as random cropping, image rotation, and flipping, the diversity of the training dataset can be increased. This enhances the model's resilience to variations encountered in the real world by preventing it from being impacted by orientation or partially obstructed images when recognizing animals.

Optimization of hyperparameters, including learning rate, batch size, and optimizer selection, has a substantial effect on the performance of a model. More than adequate selection of hyperparameters may impede convergence, result in insufficient generalization, or potentially lead to overfitting. The quality of input data is enhanced through image preprocessing techniques, including noise reduction, contrast enhancement, and normalization, which facilitate the extraction of significant features. By maintaining pixel values within a consistent range, normalization prevents numerical instability during the training process. The architecture of ResNet50V2 offers a robust framework for tasks involving animal classification. Nevertheless, further adjustments, such as the incorporation or elimination of strata, might be investigated to enhance accuracy with regard to specific animal species. Ensuring that a model strikes a balance between complexity and generalizability is crucial to prevent overfitting.

The training methodology consists of the selection of the loss function, the optimization algorithm, and the duration of training. An assortment of loss functions, including categorical cross-entropy and cross-entropy, can be utilized to highlight distinct facets of classification precision. Stochastic Gradient Descent (SGD) or Adam are optimization algorithms that govern the process by which the model updates its weights throughout the training phase. Maximizing accuracy requires finding the best combination of these parameters. Additionally, addressing data imbalances is vital. When it comes to minority classes, an unbalanced distribution of animal species can impact the precision of a model. Implementing oversampling, undersampling, or weighted sampling strategies can aid in data balancing and guarantee that the model acquires the ability to classify all species accurately. Regularization methods, including L1 and L2 regularization, serve the purpose of penalizing overly complex models to mitigate overfitting. Through the utilization of ensemble learning, precision can be enhanced by capitalizing on the benefits of numerous models and mitigating the influence of individual errors. Ongoing assessment of the model using validation datasets is critical to detect issues and identify opportunities for enhancement. Consistently augmenting the model with fresh data and refined training methodologies can yield a substantial enhancement in precision.

Related Work

Prior studies have investigated the application of ResNet50V2 in the classification of animals, with a particular focus on the impact of model architecture, data quantity, and augmentation methods on its performance. Among the areas of emphasis are the resolution of data imbalances and the optimization of hyperparameters to improve the classification precision of animal species. Forest fires are dangerous due to climate change, temperature rise, lightning, volcanoes, and humans. Detection and response matter. This study solves this problem using deep learning sub-topic transfer learning methods, achieving 97.95% classification accuracy and 99.32% transfer learning accuracy [14]. A facial image, machine vision, and deep learning model are used to identify sheep in the paper. The system achieved 95% accuracy with transfer learning on 81 Assaf breed sheep [15]. Four transfer learning ensemble CNNs are compared to state-of-the-art CNN architectures in this study. EnsenbleDVX, the best ensemble CNN, averages 97.7% accuracy [16]. Using sound and reference vocalization data from citizen scientists, this study developed a CNN pipeline to detect 54 bird species in Sonoma County, California. The optimal CNN architecture was 84.5%, and acoustic pre-training and fine-tuning improved accuracy by 10.3%. However, low-fidelity ARUs, background noise, and overlapping vocalizations occurred [17]. This paper describes a fast, accurate, fully automated chest CT scan COVID-19 detection method. Its new dataset and image processing algorithm reduce processing time and false detections. Single image classification and patient condition identification are 98.49%

accurate [18]. The study shows that weight gain and sheep maturation improve biometric identification accuracy in training and testing on 2-month-old images [19].

4. Conclusion

Five animal species were classified utilizing the ResNet50V2 deep-learning model: pandas, cats, dogs, cows, and elephants. The improved image representations of pandas and elephants contributed to the model's successful identification of these species. Methods for augmenting data were applied to improve the capability of distinguishing between classes that exhibit visual similarities. This study underscores the significance of utilizing diverse datasets when developing Deep Learning algorithms for computer vision. Such datasets guarantee an equitable distribution of classes and image quality. These findings underscore the significance of incorporating a wide range of datasets when designing Deep Learning algorithms for computer vision. Additional ramifications of this study extend to the domain of animal conservation, wherein the implementation of automatic animal recognition technology facilitates the surveillance and safeguarding of endangered or vulnerable species. In a broader sense, the application of Deep Learning models paves the way for the creation of more advanced object recognition systems in diverse visual environments, including the detection of wild animals in ecological research.

Fifty layers comprise the ResNet50V2 architecture, which also incorporates batch normalization, convolution blocks, and ReLU activation. Figure 2 shows the model's internal architecture and interconnections. When trained at an increased learning rate, the model attained a validation accuracy of 96% and a training accuracy of 98%. Following defrosting, the model underwent training at a reduced learning rate to optimize its weights and improve its robustness. With a training loss ranging from 0.01 to 0.08, the ultimate iteration of the model attained a validation accuracy of 97% and a training accuracy of 99%.

5. References

[1] Hindarto, D., 2023. Battle Models: Inception ResNet vs. Extreme Inception for Marine Fish Object Detection. *Sinkron: jurnal dan penelitian teknik informatika*, *8*(4), pp.2819-2826. DOI: https://doi.org/10.33395/sinkron.v8i4.13130.

[2] Hindarto, D., 2023. Enhancing Road Safety with Convolutional Neural Network Traffic Sign Classification. *Sinkron: jurnal dan penelitian teknik informatika*, *8*(4), pp.2810-2818. DOI: https://doi.org/10.33395/sinkron.v8i4.13124.

[3] Sze, E., Santoso, H. and Hindarto, D., 2022. Review Star Hotels Using Convolutional Neural Network. *Sinkron: jurnal dan penelitian teknik informatika*, *7*(4), pp.2469-2477. DOI: https://doi.org/10.33395/sinkron.v7i4.11836.

[4] Iswahyudi, I., Hindarto, D. and Santoso, H., 2023. PyTorch Deep Learning for Food Image Classification with Food Dataset. *Sinkron: jurnal dan penelitian teknik informatika*, *8*(4), pp.2651-2661. DOI: https://doi.org/10.33395/sinkron.v8i4.12987.

[5] Hindarto, D. and Santoso, H., 2022. Performance Comparison of Supervised Learning Using Non-Neural Network and Neural Network. *Jurnal Nasional Pendidikan Teknik Informatika: JANAPATI*, *11*(1), pp.49-62. DOI: https://doi.org/10.23887/janapati.v11i1.40768.

[6] Hindarto, D., 2023. Comparison of Detection with Transfer Learning Architecture RestNet18, RestNet50, RestNet101 on Corn Leaf Disease. *Jurnal Teknologi Informasi Universitas Lambung Mangkurat (JTIULM)*, *8*(2), pp.41-48. DOI: https://doi.org/10.20527/jtiulm.v8i2.174.

[7] Hindarto, D., 2023. Exploring YOLOv8 Pretrain for Real-Time Detection of Indonesian Native Fish Species. *Sinkron: jurnal dan penelitian teknik informatika*, *8*(4), pp.2776-2785. DOI: https://doi.org/10.33395/sinkron.v8i4.13100.

[8] Du, L., Lu, Z. and Li, D., 2022. Broodstock breeding behaviour recognition based on Resnet50-LSTM with CBAM attention mechanism. *Computers and Electronics in Agriculture*, *202*, p.107404. DOI: https://doi.org/10.1016/j.compag.2022.107404.

[9] Boulila, W., Alzahem, A., Koubaa, A., Benjdira, B. and Ammar, A., 2023. Early detection of red palm weevil infestations using deep learning classification of acoustic signals. *Computers and Electronics in Agriculture*, *212*, p.108154. DOI: https://doi.org/10.1016/j.compag.2023.108154.

[10] Panthakkan, A., Anzar, S.M., Jamal, S. and Mansoor, W., 2022. Concatenated Xception-ResNet50—A novel hybrid approach for accurate skin cancer prediction. *Computers in Biology and Medicine*, *150*, p.106170. DOI: https://doi.org/10.1016/j.compbiomed.2022.106170.

[11] Sharma, A.K., Nandal, A., Dhaka, A., Zhou, L., Alhudhaif, A., Alenezi, F. and Polat, K., 2023. Brain tumor classification using the modified ResNet50 model based on transfer learning. *Biomedical Signal Processing and Control*, *86*, p.105299. DOI: https://doi.org/10.1016/j.bspc.2023.105299.

[12] Sharma, A.K., Nandal, A., Dhaka, A., Polat, K., Alwadie, R., Alenezi, F. and Alhudhaif, A., 2023. HOG transformation based feature extraction framework in modified Resnet50 model for brain tumor detection. *Biomedical Signal Processing and Control*, *84*, p.104737. DOI: https://doi.org/10.1016/j.bspc.2023.104737.

[13] Rahimzadeh, M. and Attar, A., 2020. A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2. *Informatics in medicine unlocked*, *19*, p.100360. DOI: https://doi.org/10.1016/j.imu.2020.100360.

[14] Reis, H.C. and Turk, V., 2023. Detection of forest fire using deep convolutional neural networks with transfer learning approach. *Applied Soft Computing*, *143*, p.110362. DOI: https://doi.org/10.1016/j.asoc.2023.110362.

[15] Hitelman, A., Edan, Y., Godo, A., Berenstein, R., Lepar, J. and Halachmi, I., 2022. Biometric identification of sheep via a machine-vision system. *Computers and Electronics in Agriculture*, *194*, p.106713. DOI: https://doi.org/10.1016/j.compag.2022.106713.

[16] de Jesus Silva, L.F., Cortes, O.A.C. and Diniz, J.O.B., 2023. A novel ensemble CNN model for COVID-19 classification in computerized tomography scans. *Results in Control and Optimization*, *11*, p.100215. DOI: https://doi.org/10.1016/j.rico.2023.100215.

[17] Clark, M.L., Salas, L., Baligar, S., Quinn, C.A., Snyder, R.L., Leland, D., Schackwitz, W., Goetz, S.J. and Newsam, S., 2023. The effect of soundscape composition on bird vocalization classification in a citizen science biodiversity monitoring project. *Ecological Informatics*, *75*, p.102065. DOI: https://doi.org/10.1016/j.ecoinf.2023.102065.

[18] Rahimzadeh, M., Attar, A. and Sakhaei, S.M., 2021. A fully automated deep learning-based network for detecting COVID-19 from a new and large lung CT scan dataset. *Biomedical Signal Processing and Control*, *68*, p.102588. DOI: https://doi.org/10.1016/j.bspc.2021.102588.

[19] Hitelman, A., Edan, Y., Godo, A., Berenstein, R., Lepar, J. and Halachmi, I., 2022. The effect of age on young sheep biometric identification. *Animal*, *16*(2), p.100452. DOI: https://doi.org/10.1016/j.animal.2021.100452