Endangered Animal Recognition System

Lim Zhe Yu, Nicole Ha Yah Yie, Ooi Weishan, Tan Juan Kai

0- Abstract

The United Nations' Sustainable Development Goals (SDGs) encompass 17 global initiatives designed to tackle global challenges and promote sustainable development across a variety of sectors. Among these, SDG 15, "Life on Land," is dedicated to the preservation of wildlife species and their habitats, aiming to maintain balanced biodiversity within our ecosystems. This goal is crucial as it protects vital resources for human survival, such as food and oxygen. In recent years, Computational Intelligence (CI) has surfaced as a potent instrument in achieving this goal. This paper will examine recent CI methodologies from 2018 to 2023, with a focus on fuzzy logic, machine learning, and deep learning techniques. To be more specific, this research will be emphasized on developing endangered animal recognition system using both machine learning approach and deep learning approach. Moreover, concept of feature aggregation will be introduced by combining features extracted by pretrained model namely EfficientNetB2 and VGG16, then combined together for classification purposes. On top of that, to effectively determine the suitable machine learning model and its optimal parameter, Lazy Predict will be utilized in this research together with Random Search. Experiments had been conducted and concluded that MLP having the best performance among all the classification model, with an accuracy of 92.50% on Dataset 1, 89.58% on Dataset 2, with average performance of 91.04%.

Keywords: SDG; Life on Land; Machine Learning; Deep Learning; Endangered Animal Recognition; Feature Aggregation; EfficientNetB2; VGG16; Random Search; Lazy Predict

1- Introduction

In the field of Computer Intelligence (CI) indicates as the convergence of artificial intelligence, machine learning and computational modeling. The primal objective is to provide a robust toolkit that could adapt to real like situations and addresses real world problems. The discourse surrounding Sustainable Development Goals (SDGs) has gained substantial momentum owing to its global significance and the concerted commitment of nations worldwide towards its realization. SDG are sets of target and indicator consists of 17 main pillars in order to raise and combat rising global issues such as ending poverty and hunger, promote gender equality, preserving environmental diversity, addressing of economic growth and others by looking at 5 main dimensions which are people, planet, prosperity, peace and partnership [1]. SDG 15, aptly named as "Life on Land", this goal indicates as a global initiative underscoring the pivotal role of terrestrial ecosystems. A vast array of cultural, spiritual, and economic values is represented by these ecosystems. Not only are they necessary for human survival, but they also produce more than half of the global GDP. Sadly, pollution, biodiversity loss, and climate change together represent a triple threat to the world. The global trends of rising deforestation, land degradation, and species extinction pose a major threat to the planet and its inhabitants. Though there has been some progress in protected areas, sustainable forest management, measuring natural resources, and national biodiversity values, these advancements have shown to be limited [2]. Thus, the applications of computational intelligence (CI) as a potent tool for achieving SDG 15 indicated as a vital tool for enhancing efficiency and productivity in certain domains.

In order to evaluate intricate environmental data and support the preservation and sustainable use of terrestrial ecosystems, CI makes use of sophisticated algorithms and computational models. Numerous techniques are available for CI applications, such as deep learning, neural networks, machine learning, fuzzy systems, and so on. A type of many-valued logic known as fuzzy logic is skilled at modeling ambiguity and uncertainty, which are frequently present in ecological data. Machine learning is a versatile approach that can handle both structured and unstructured data since it utilizes statistical approaches for autonomous learning and decision rule building. Deep learning is a subset of machine learning that offers improved dependability, particularly in noisy data environments, by using numerous layers of artificial neural networks to identify complex patterns in huge datasets.

Upon to be more specific and delving into the topic, this project will be focusing on developing an endangered animal recognition system with machine learning and deep learning approaches. In the context of applying both machine learning and deep learning approaches, 4 machine learning models and 4 deep learning models, in total of 8 model will be included in this project to compare the performances across two online datasets from Kaggle. The selection of the machine learning models and its optimal hyperparameters are determined by applying Lazy Predict with Random Search experiments. Besides, the selection of the deep learning models is determined by the commonly applied deep learning models to investigate its robustness on image classifications. In order to perform consistent feature extraction from the animal images, this project will be applying two pretrained model namely EfficientNetB2 and VGG16 to extract meaningful features from the images and then combine together to feed into all the models to perform classification. This comprehensive approach ensures that the review not only highlights the successes but also addresses the complexities and challenges in this field.

2- Related Work

2-1- Fuzzy System

Fuzzy systems take into account the inherent uncertainty in human reasoning by permitting components to belong to a set with a certain degree of certainty. Fuzzy systems enable the modeling of common sense and are based on fuzzy sets and logic. In this section, a brief overview of animal detection systems using fuzzy system will be presented.

2-2-1- Animal Detection Using Fuzzy System

Fuzzy systems have a wide range of applications, including animal detection systems. These systems use certain rules or formulas to recognize various endangered species in image or video format [3], [4]. For instance, research [3] presented an IF-FUZ-based system for animal identification and recognition that makes use of fuzzy logic and thermal imaging of the animals. Using invariant properties or input such as Zernike moments, skeleton routes, local binary patterns, and form features, the fuzzy system extracts data from thermal pictures. Based on a series of fuzzy rules and membership functions, these features are then merged and fed into a fuzzy interference system, which divides the animals into 12 distinct classes with 3 main categories.

In addition, a different study [4] proposed a next-generation roadside animal detection system (NG RADS) that estimates the various degrees of harm caused by animals on the roadway using fuzzy rule-based algorithms. The four input indications used in the scenario and threat assessment phase used are mainly physical distance, stopping distance, speed magnitude, and the distance from the animal to the roadway. These signs are utilized to assess the degree of threat posed by the animal and to gauge the geographical and temporal interaction between it and the vehicle.

Table 1. Table of Summary for Applications using Fuzzy System.

Domain	Reference	Methodology	Performance
Animal Detection	[3]	**Image Preprocessing:** Gamma correction, Histogram equalization, Gradient-based guided edge-aware smoothing filter **Threshold segmentation to combine of 4 features**: Zernike, skeleton path, local binary pattern and shape features (use fusion of SMM & MAM) The features are then fed into a fuzzy inference system that classifies the animals into 12 classes and three categories based on a set of fuzzy rules and membership functions.	Average Accuracy: 97%
	[4]	**Fuzzy Rule Based Algorithm:** Determination of input fuzzy set - 2 input fuzzy sets for each indicator Determination of membership function for input fuzzy set – “trapezoid” membership function Determination of boundary values for membership function Determination of output fuzzy set and fuzzy rule	N/A

2-2- Machine Learning

Machine learning (ML), a branch of computer science and artificial intelligence (AI), uses data and algorithms to gradually increase the accuracy of a model, simulating human learning. In this section, a brief overview of animal classification system and animal prediction system using various machine learning algorithms will be presented.

2-2-1- Animal Classification System using Machine Learning

Conventional ML methods have been applied to the classification of animals in different species and environments in earlier research. Researchers have investigated traditional algorithms including Support Vector Machines (SVM), Logistic Regression, K-Nearest Neighbors (KNN), and Decision Trees, with a particular focus on three endangered species. Remarkably, the accuracy of the ensemble classification using Random Forest (RF) was only 74.2%, but the accuracy of the Logistic Regression model was a noteworthy 93.1% [5]. A comparison between RF and Multi-Layer Perceptron (MLP) in the context of Brazilian parrot species demonstrated that RF might achieve up to 99% accuracy in certain instances [6]. SVM-based bird species classification yielded 98% training and validation accuracy [7]. In addition, SVM was employed by [8] to study various transfer learning (TL) techniques for animal categorization, with a training time of 57.15 seconds and an average accuracy of 97.5%. For the best result of 79.67% accuracy in classifying wild boar, a combination of texture analysis techniques and Linear Discriminant Analysis (LDA) was employed [9].

2-2-2- Animal Prediction System using Machine Learning

The forecast of endangered animals and population monitoring for the upcoming timeframe have made extensive use of AI and ML algorithms. In recent studies, researchers [10] have attempted to use decision trees, random forests, and linear regression models in endangered species conservation by projecting the total population of endangered animals for the ensuing ten years using time series data. Of the three proposed models, linear regression achieved the highest score of 94.3%.

Additionally, an alternate XGBoost model based on machine learning and easily accessible data was presented as an automated evaluation approach to determine the extinction risk of reptile species. In order to capture significant features prior to training, the suggested methodology made use of phylogenetic and spatial eigenvector maps to account for spatial and phylogenetic autocorrelation. The suggested classifier is believed to have extremely high accuracy, having attained 90% in categorizing species as threatened or nonthreatened and 84% in predicting particular extinction risk categories [11].

Table 2. Table of Summary for Applications using Machine Learning.

Domain	References	Methodology	Performance
Animal Classification	[5]	**Classical Classification:** Logistic Regression **Ensemble Classification:** Random Forest	**Accuracy:** Logistic Regression: 93.1% Random Forest: 74.2%
	[6]	Random Forest	Accuracy: up to 99%
	[7]	SVM	Training accuracy: 98% Validation accuracy: 98%
	[8]	VGG16 + SVM	Average accuracy: 97.5% Training time: 57.15 sec
	[9]	**3 Classes Scenario:** LBP+FD with LDA (single method) LBP+FD, RNN and lacunarity (local) with LDA (combination method) **2 Classes Scenario:** GLDM with LDA (single method) GLDM, Gabor filters, LBP+FD, and RNN with LDA (combination method)	Single Accuracy: 64.89% Combination Accuracy: 68.56% Single Accuracy: 73.44% Combination Accuracy: 79.67%
Animal Prediction	[10]	Linear Regression	Accuracy: Linear Regression: 94.3%

[11]	XGBoost	Classify species as threatened/nonthreatened: 90% acc Predicting specific extinction risk categories: 84% acc

2-3- Deep Learning

Deep learning (DL) is the subset of ML methods based on artificial neural networks and representation learning. Computers can be trained to interpret information similarly to the human brain through a technique called deep learning. In this section, a brief overview of animal detection system, animal prediction system, animal classification system and animal poaching system using deep learning algorithms will be presented.

2-3-1- Animal Detection System Using Deep Learning

Deep learning has been popular in recent years, particularly in the domain of computer vision. CNN is commonly used in creating the model for animal detection system in monitoring agricultural lands [12]. A result of 86% of IOU and 99% average F1 score were achieved on IWildCam datasets by using CNN, in which the depth-wise separable convolution layers were used for feature extraction and object detector and fully connected layers for classification [13]. In the realm of tiger detection, [14] proposed a method that surpassed the other state-of-the-art models. It has attained a result of 60.8% mean average precision (mAP) by using Fast R-CNN for object detection and softmax function as classifier. Similarly, [15] achieved a 81% accuracy on a publicly available dataset on Kaggle by employing a CNN with fully connected layers and softmax activation function.

Using pre-trained models is a popular approach in objection detection. You Only Look Once (YOLO) [16], [17], [18], [19], ResNet [18], [20], [21], VGG [20], [21], MobileNet [18], [22], [23], and Xception model [24] are some of the pre-trained models that are frequently used. To decrease training time while maintaining a high-performing model, these pretrained models are typically adjusted to meet the needs of relevant research fields. To improve the fusion capabilities and feature enhancement of YOLOv5, [16] added Weighted Bi-directional Feature Pyramid Network (BiFPN) and Efficient Channel Attention (ECA) to the model, achieving an accuracy of 95.5% for wildlife detection. Conversely, [17] performed transfer learning on the YOLOv5 model, utilizing pre-trained weights from MS-COCO datasets, achieving a 94% mAP and an average recall of 90%. Additionally, [25] discovered that YOLOv5 outperformed ResNet and HRNet32 in animal detection and classification on camera trap images, with an average accuracy on video frames reaching 89.3%. On a different note, [19] achieved a 91% mAP detection rate by modifying the head, neck, and backbone of YOLOv5 during the feature extraction stage. Similarly, [26] improved model stability and attained a 95% mAP identification rate by introducing alterations to the network's feature extraction components and employed the Gated Linear Unit (GLOU) loss function in the classifier.

MobileNet-SSD V2 coco and MobileNet-SSD V2 320 x 320 are proven to be effective for object detection and image processing in ecological surveillance fields. A fully connected layer at CNN can be used as a classifier to achieve 95% accuracy, and the model is also shown to be robust by achieving an 85% accuracy rate even in cases where the image is slightly blurred [22]. [18] used the pretrained weight of ImageNet to merge two pretrained models, YOLO and MobileNet, in recognizing endangered parrot species with a result of mAP of 86.76%. [23] has used a mix of CNN and MobileNet to train a detection model that was tested on three different datasets: ARE, Animals Detection Images Dataset, and Google Open Images V6 datasets. The model achieved accuracies of 92%, 99.6%, and 95.6%, respectively.

In order to recognise and distinguish between different types of animals, such as snakes, lizards, and others, [20] integrated the pretrained model of ResNet50 as an object detection model and later fully connected with two dense layers and SoftMax activation layer as classifier. Using the camera trap dataset, the model accuracy achieved an overall percentage of 86% for the multiclass experiment and an average accuracy of 92% for the binary class experiment. In the work of [24], the Xception model with 36 convolutional layers and depth-wise separable convolutions for object detection is used, along with a custom network with four fine-tuning layers, which achieved 96% accuracy and 96.1% F1-score in detecting and classifying wild cats in ecological systems. For Pangolin detection and classification, [21] utilised the pre-trained VGG16 model for classification and Fast R-CNN ResNet 101 with MS COCO pre-trained weights for object detection tasks. This model achieved outstanding accuracy of 96.38%, a specificity of 97.55%, and an F1-score of 96.36%, and a mAP of 93.41%.

On the other hand, there are other deep learning methods for classifying and detecting animals other than CNN. For example, [27] presented a KI-CLIP technique that blends textual expert knowledge and image data based on the pre-trained CLIP model for few-shot and incremental recognition of uncommon and endangered wildlife, without using any CNN techniques. The performance of KI-CLIP is assessed on 12 datasets in various few-shot settings, comparison with four baseline models, ablation experiments and case studies. Then, using Sync Dreamer and NeRF, [28] had created an animal detection model. SyncDeamer is a diffusion model that creates multiview-consistent images from a single-view image of any object. In the meantime, a 3D representation-generating model called NeRF is able to learn a continuous 5D function that translates a 2D viewing direction and a 3D position into a volume density and a view-dependent colour. Conversely, [29] created a FRW-ACA detection model through the use of ensemble network techniques, and they were able to achieve a 74.91% mAP detection rate.

2-3-2- Animal Prediction System Using Deep Learning

CNN, in particular, are a deep learning approach that has shown useful in assessing dynamic traits, most notably in forecasting the susceptibility of endangered species to climate change, as evidenced by recent research [30], [31], [32]. In order to provide meaningful information on endangered species, researchers [30] conducted research in this field and used CNN for classification tasks based on photographs of such species. Furthermore, by utilizing historical data, conventional machine learning methods like logistic regression and linear regression were integrated to predict future changes in temperature as well as the number of endangered animals. The study achieved promising results, with a detection accuracy of 96.6% for identifying images containing animals and 90.4% accuracy in identifying the three most prevalent species among wild animals captured in South-central Victoria, Australia.

Additionally, another study employed CNN models to forecast the likelihood of species extinction based on the International Union for the Conservation of Nature's (IUCN) Red List criteria. CNN-class, BNN-class, and NN-reg models were used for classification tasks, while a 2D CNN model was used for feature extraction. With an accuracy of 0.60 for each of the five Red List categories and 0.81 for differentiating between potentially threatened and non-threatened species, the NN-class model yielded the best overall results [31].

2-3-3- Animal Classification System Using Deep Learning

In the several existing works like [15], [33], [34], [35], [36], [37], [38], CNN were widely applied on animal classification with deep learning methods, where various CNN architectures were applied. For instance, in the work in [33], CNN was utilized to identify wild species in Texas and monkey species in work [15] while achieving validation accuracies of 60% and 81% respectively. In work [34] and [36], both works have utilized InceptionResNetV2 in wildlife classifications and endemic bird classifications. [34] has achieved 91% mean Average Precision (mAP) for testing, [36] successfully classify endemic bird species with an accuracy of 98.39%. Further, Furthermore, the ResNet-18 model was used to identify animals with camera trap and achieved 96.8% species model accuracy and 97.3% empty animal model accuracy. For the classification of endangered parrots, a Multilayer Perceptron (MLP) has achieved accuracies ranging from 89.83% to 98.17% [39], and a CNN architecture named NASNetMobile achieved an accuracy of 94.13%. Additionally, a unique methodology that involved the use of a fine-tuned MegaDetector model from [37], combined with Faster R-CNN with an Inception-ResNet-v2 as the backbone for wildlife classification. It has achieved with mAP performance of 74% on the test set, 97% mAP at IoU and 89.24% mAP at IoU.

2-3-4- Animal Poaching Using Deep Learning

In the research of [40], the work introduced an automated weapon detection system that was deployed in forest areas that aims to prevent poaching and protect endangered animals. This system utilized YOLOv5 which is a well-known and widely applied object detection algorithm on a thermal infrared dataset by ASL labs. However, the paper did not provide any specific performance metrics for the system. On the other hand, another study from [41], the focuses were on detecting and identifying poachers and wildlife using thermal cameras that are attached to drones. The researchers adopted a two-layered CNN to classify the objects within the bounding boxes as human or animals. It also implemented a centroid tracking algorithm for tracking objects across frames. Overall, the system demonstrated by [40] has an impressive identification accuracy. It achieved 100% accuracy for humans, 98.33% for cats, and 99.67% for horses when tested on the Thermal Infrared Dataset by ASL Labs. The system also showed high overall tracking accuracy, with 90.93% for humans, 94.55% for cats, and 95.88% for horses. The study in [41] emphasized the effectiveness of low-cost thermal cameras and contemporary computer vision techniques in detecting, differentiating and tracking humans and animals. This showcases its potential for detecting poachers and counting wildlife, especially in game reserve settings during nighttime conditions. These findings highlight the importance of deep learning applications in creating advanced and efficient systems for wildlife conservation and protection.

Table 3. Table of Summary for Applications using Deep Learning.

Domain	References	Methodology	Performance
	[16]	**Feature Extraction/ Object Detection**	95.5% accuracy

		ECA and BiFPN applied on YOLOv5 for better fusion capabilities and enhancement of feature **Classifier:** Depth-separable convolutional layer	
	[22]	**Feature Extraction/ Object Detection:** MobileNet-SSD V2 coco and MobileNet-SSD V2 320x320 **Classifier:** CNN	Overall performance accuracy rate 95% Accuracy rate of over 85% for slightly blurred images IOU: 86% for detection performance
	[13]	**Feature Extraction/ Object Detection** Depth-wise separable convolution layers (CNN) **Classifier** Fully connected layers (CNN)	Precision, recall and F1 score of tigers is obtained as 1.00, 0.99, 0.99 respectively. Precision, recall and F1 score of raccoon is obtained as 1.00, 1.00, 1.00 respectively. Precision, recall and F1 score of cheetahs is obtained as 0.99, 1.00, 1.00 respectively.
Animal Detection	[25]	**Feature extraction, object detection and classifier** YOLOv5, ResNet50 & ResNet101 + FCOS, HRNet32 + Cascade R-CNN	**Best performance:** YOLOv5: average accuracy of 89.3% on video classification
	[17]	**Feature extraction, object detection and classifier** YOLOv5 model with transfer learning, using a pre-trained model on MS-COCO dataset	mAP 0.5: 94% Average recall: 90%
	[12]	**Feature extraction, object detection and classifier** CNN	N/A
	[14]	**Feature extraction / Object Detection** Faster R-CNN **Classifier** Fully connected layer using softmax activation function	mAP: 60.8%
	[21]	**Feature extraction / Object Detection** Faster R-CNN ResNet50, Faster R-CNN ResNet101, Faster R-CNN Inception v2 **Classifier** VGG16, EfficientNetB0, EfficientNetB1, DenseNet121, DenseNet201	Classification of With Pangolin vs No Pangolin: (best performance) VGG16: 96.38% accuracy, 97.55% specificity, 96.36 F1-score Detection: Faster R-CNN ResNet101 (with MS COCO pretrained weight): 93.41% (mAP)
	[27]	4 modules: Specialist knowledge extraction, pre-trained CLIP model, self-attention multilayer perceptron (SA-MLP), incremental learning mechanism 3 stages: Few-shot/zero-shot learning, few-shot/zero-shot inference, incremental learning	Accuracy: 90.76%

[18]	MobileNet + YOLO	86.79% of mean average precision
[20]	**Feature extraction/ Object Detection:** VGG16, ResNet50 **Classifier:** Fully Connected layer with 2 dense layers and SoftMax activation layer	**Best performance** 86% accuracy for multiclass experiment using ResNet50 92% average accuracy on binary class experiment using ResNet50
[24]	**Feature extraction/ Object Detection:** Xception Model (36 convolutional layers and uses depth-wise separable convolutions) **Classifier:** Custom network that consists of four fine-tuning layers	Accuracy: 96%, Sensitivity: 96%, Specificity: 99.4%, Precision: 96.2%, F1 score: 96.1%
[23]	MobileNet and Single Shot MultiBox Detector (SSD)	**ARE dataset:** Accuracy: 92%, Precision is 93% Recall: 91% F1-score: 95% **Animals Detection Images Dataset:** Accuracy: 99.6% Precision 99% Recall: 98% F1-score: 99% **Google Open Images V6 dataset** Accuracy: 95.6% Precision 95% Recall: 96% F1-score: 96%
[28]	SyncDreamer – Generate novel views of endangered species NeRF - Create 3D representations	N/A
[15]	**Feature extraction**: CNN **Classifier**: Fully connected layers wirh softmax activation function	Accuracy: 81%
[29]	**Feature Extraction:** Feature reweighting module **Classifier:** Ensemble learning of relation network with MSE loss function	74.91 mAP using proposed FRW-ACA method
[19]	**Feature Extraction:**	mAP: 91%

		YOLOv5: Backbone - Cross Stage Partial DenseNet (CSPDenseNet) Neck - The Path Aggregation Network (PA-Net) utilizes bottom-up path augmentation to improve and reduce the feature pyramid as well as correct low-level signal localization. Head - applied anchor boxes on features > generate output features	
	[26]	**Classifier (Transfer learning):** YOLOv5 model **YOLOv5** **Backbone network:** Consists of a Focus module, a CSP module, and an SPP module. **Neck network:** A set of feature aggregation layers that build a Feature Pyramid Network (FPN) to enhance the feature fusion and detection of objects with different sizes and scales. **Detection network:** 3 detection layers, each receiving a feature map with different resolutions. **Classifie**r: GIoU loss function	mAP: 95%
	[30]	**CNN**	Images containing animals: 96.6% Three most common wild animals: 90.4%
Animal Prediction	[31]	**2D-CNN**	nn-class:0.6, nn-reg:0.54, bnn: 0.6
	[32]	**MLP**	MLP with node2vec + doc2vec: F1 score = 0.885, MCC = 0.583, ROC-AUC = 0.873
	[33]	CNN	Training accuracy: 69% Validation accuracy: 60%
	[34]	InceptionResNetV2	mAP: 91%
	[35]	ResNet-18	Species model accuracy 96.8% Empty-animal model accuracy 97.3%
	[15]	CNN	Accuracy: 81%
Animal Classification	[39]	MLP	Accuracy ranging from 89.83% to 98.17%, depending on the scenario
	[36]	Inception-resNet-v2	**Classification of endemic bird species** Accuracy: 98.39% Precision: 98.49% Recall: 97.50% F1-score: 97.90%

			Detection of birds among different object categories: 100%
	[37]	Fine-tuned MegaDetector model (composed of Faster R-CNN with Inception-ResNet-v2 backbone)	mAP: 74% mAP: 97% (IoU=0.50) mAP: 89.24% (IoU=0.754)
	[38]	NASNetMobile	Accuracy: 94.125% F1-score: 95%
	[40]	YOLOv5	N/A
Animal Poaching	[41]	Two-layered convolutional neural network (CNN) to classify the objects within the bounding boxes as humans or animals. Centroid tracking algorithm to assign unique IDs to the detected objects and track them across frames.	**Overall identification accuracy** Human: 100% Cat: 98.33% Horse: 99.67%

3- Proposed Method

The proposed method is a Multilayer Perceptron (MLP) based architecture. The pipeline of the proposed MLP method is as depicted in Figure 1. Image resizing and data augmentation were used to pre-process the image data. The pre-trained model VGG16 and EfficientNetB2 were used to extract the features of the pre-processed images. The features extracted from both VGG16 and EfficientNetB2 were then combined to form combined features. These combined features were then passed to MLP for model training and classification of the endangered animal. The optimal value of hyperparameters of MLP was determined using random search strategy. Lastly, the results were evaluated on 4 performance metrics, i.e. accuracy (AC), precision (PR), recall (RC), and F1-score (F1).

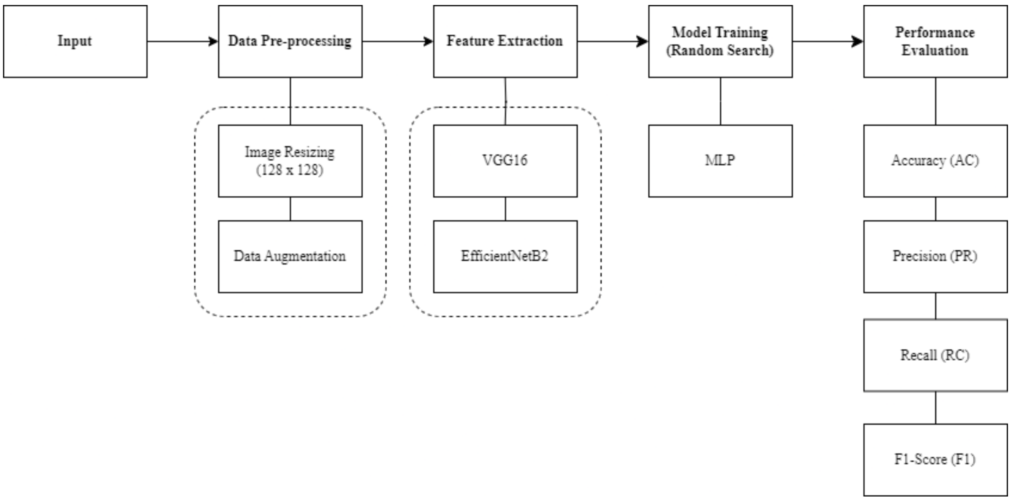

Figure 1. Pipeline of the proposed method.

3-1- Preprocessing

In the data preprocessing step, the dataset was first resized to 128 x 128 pixels. The dataset is then split into training and testing sets using an 80-20 split. To increase the diversity and number of training data, data augmentation is applied on the dataset. This involved rotation, width and height shifts, shear, zoom, and horizontal flipping. Each original image in the dataset generated 5 variations of augmented images as shown in Figure 2. Common data preprocessing techniques such as image grayscaling was not employed in this experiment, the RGB information was retained as colour information is an important feature for animal classification.

Figure 2. Sample output of 5 augmented images.

3-2- Feature Extraction

In order to compare the performance of all the models under the same condition, this project has applied two pre-trained model which are EfficientNetB2 and VGG16 to extract meaningful features from the images and then combine the extracted features from both pre-trained models to feed into all the models to perform animal classification. For EfficientNetB2, which is a variation of CNN, it extracts features from the image and feeds them into a deep neural network, which generates a sequence of probabilities representing the likelihood that the image falls into each of those classes. The highest probability output, according to the probability distribution is the model's prediction. Apart from EfficientNetB2, another CNN variant is VGG16 is proposed for feature extraction. VGG16 has been known on its effective ness on extracting features from images by breaking down the image into smaller sections and analyzes each one of them. The model is composed of multiple convolutional layers which are followed by one or more dense layers that are fully completely connected. After the images of the dataset are loaded into the VGG16 model, they are expanded from a three-dimensional array to a four-dimensional array with dimensions. Next, it is transformed to a NumPy array of pixel data and then the features are ready to be obtained.

Both models are adopted by removing the last fully connected layer and the extracted features are then saved to pickle files. The extracted features from both models are combined as one pickle file to obtain the complementary features as both models might obtain different features based on different perspectives. Thus, it is assumed to increase the performance and robustness of the model. Both models applied the same approach to breaking down the image into smaller sections and analyzing each one. The other reason for implementing these two models is due to their efficiency and accuracy on a variety of scales, which provides a pivotal help to this project.

3-3- Multi Layer Perceptron (MLP)

MLP is a type of Artificial Neural network, in which the nodes inside the neural network represent the human neurons. A basic MLP network consists of 3 layers, which are input layer, output layer, and hidden layers in between. The number of neurons in each layer should have been identified. The input layer is automatically set by the algorithm, by selecting the number of features as the neurons in the first layer.

During training, the MLP adjusts the connections between neurons in order to increase its predictive accuracy. By comparing its predictions to actual outputs, the MLP updates these connections to minimize the disparity between predicted and real values. Hidden layers assist it in learning complicated patterns in data. The MLP may adapt to different types of issues by altering the number of neurons and layers and applying different mathematical functions. After it has been trained, the MLP can be used to predict outcomes for new data. predict outcomes for new data.

Figure 3 depicts the network architecture of the proposed MLP. The proposed MLP architecture is a simple and light model with input layer, output layer, and only 1 hidden layer with 100 neurons. It learns the pattern of the features fetched from the input layer. Each neuron in the hidden layer computes a weighted sum of its input features, and these weighted sums collectively capture complex patterns in the input data. Then it utilised the logistic activation function to get the predicted animal output. The rationale behind the design of this architecture was through experimenting using random searches to find the best parameters by exploring different combinations of hyperparameters in obtaining good results as shown in Table 4. The logistic activation function is represented by Equation (1).

$$f(x) = \frac{1}{1 + e^{-x}} \tag{1}$$

where *x* represents the net input to the logistic activation function, which is the weighted sum of the inputs in the hidden layer before passing through the logistic activation function.

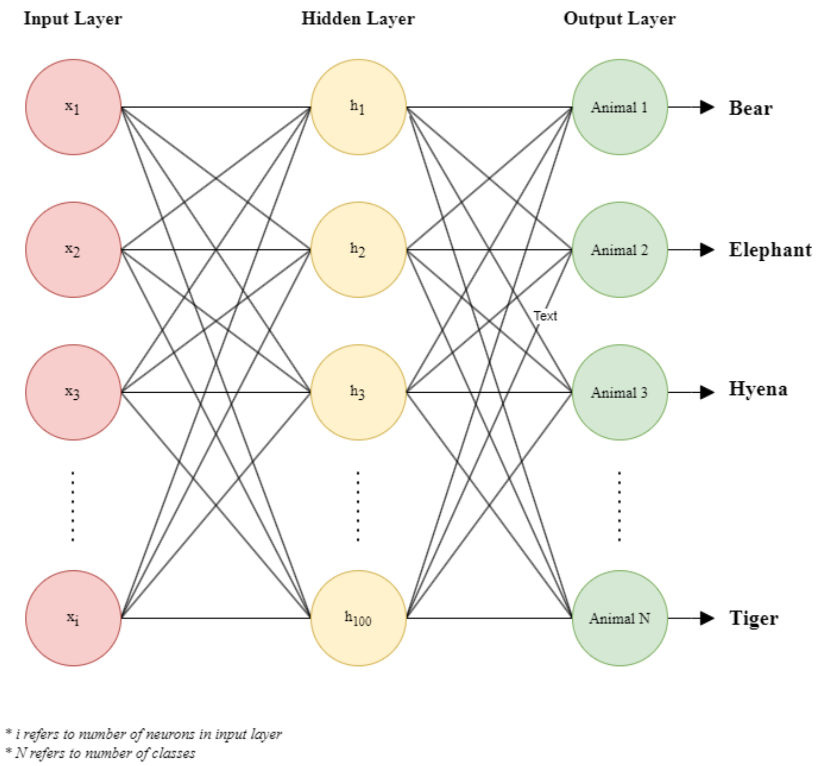

Figure 3. **Architecture of the proposed MLP.**

Table 4. **Parameter settings of the proposed MLP.**

Parameters	Value
hidden_layer_sizes	(100,)
max_iter	1000
learning_rate_init	0.01
alpha	0.01
activation	logistic

4- Experimental Results and Analysis

4-1- Datasets

Under Google LLC, Kaggle is a platform for data science competitions and an online community for data scientists and machine learning professionals. Kaggle provides a platform for quality data exploration, analysis, and sharing for practitioners in machine learning and data science. Both the datasets used in this experiment were downloaded from Kaggle, which include the Animal 151 [42] and the Animal Image Dataset (90 Different Animals) [43] . In this experiment, Animal 151 will be referred to as Dataset 1 while Animal Image Dataset (90 Different Animals) will be referred to as Dataset 2.

First and foremost, Dataset 1 includes 151 different animal species with up to 60 images per species. However, certain classes are represented by only 30 or fewer images. Each image in the dataset has 224 × 224 pixels and was saved in the JPG format. The focus of the experiment is on developing an "Endangered Animal Recognition System," so some of the 151 species that were initially listed as endangered animals must be filtered out. Consequently, 12 classes were chosen as a subset for additional examination and model training. The following 12 classes are home to endangered species: lemur, okapi, orangutan, panda, puma, red panda, cheetah, elephant, iguana, jaguar, Komodo dragon, and tapirus. The number of images available for each of these classes is as follows: cheetah (29 images), elephant (31 images), iguana (31 images), jaguar (47 images), Komodo dragon (33 images), lemur (41 images), okapi (37 images), orangutan (30 images), panda (31 images), puma (28 images), red panda (31 images), and tapirus (29 images).

Similarly, Dataset 2 was assembled from Google Images and consists of 5400 images with a wide variety of animals from 90 different classes or categories. All of the images in the dataset were saved in JPG format, which varies in terms of dimensions. As aforementioned, the focus of the experiment is on developing an "Endangered Animal Recognition System," thus a filtering procedure was used to separate and concentrate on these particular animals. Subsequently, a subset of 8 classes was selected for further analysis. These classes of animals comprise tigers, bears, pandas, okapi, hyenas, orangutans, elephants, and rhinoceroses. There are 60 images in total that represent each of these classes.

4-2- Experiment Settings

In order to compare and contrast the 8 models that are used for classification (4 ML and 4 DL models) by preventing any unwanted changes in variables, various performance evaluation metrics are available to determine if the models are giving out any promising results. Thus, four evaluation metrics are decided to use in this project, which are accuracy, precision, recall and f1-score. For all these evaluation metrics, it has different expressions on True Positive (TP), True Negatives (TN), False Positives (FP) and False Negatives (FN). Accuracy indicates the ratio of correctly predicted observations to the total observations, and it measures the overall correctness of the model. For precision, it is the ratio of the correctly predicted TP which commonly referred as positive observations to the total predicted positives (TP + FP). For recall, it refers to the ratio of correctly predicted positive observations which are the TP to all the actual positives. The F1-score, it indicates as the average grade on the performance of the model as it finds the balance between precision and recall. This is particularly useful when the class distribution is imbalanced. For train-test split on the dataset, 80% from each dataset were retrieved as training sets while the remaining 20% are for testing sets from each dataset. The remaining 20% are served as independent data samples for the models to evaluate their performance. The equations for all the mentioned evaluation metrics are depicted in Equations (2) to (5).

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{2}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{3}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{4}$$

$$\text{F1} - \text{Score} = \frac{2}{\frac{1}{\text{P}} + \frac{1}{\text{R}}} \tag{5}$$

4-3- Implementation Details

Random search with 10 iteraions and 3-fold cross validation is performed to identify the optimal combination of hyperparameters for the MLP model. Subsequently, the best parameters identified, including the size of hidden layers = (100,), activation function = logistic, learning rate = 0.01, maximum number of iterations = 1000 and alpha = 0.01 are then utilized for making predictions on both Dataset 1 and Dataset 2.

The performance of the MLP model is compared with other methods namely Linear SVM, Calibrated Classifier – Linear SVM, Passive Aggressive, Peceptron, CNN, DNN, LSTM and GRU. Similarly, random search is performed to identify the optimal combination of hyperparameters for Linear SVM, Passive Aggressive and Perceptron. The optimal parameters identified for the Linear SVM model is not only used on the model itself but it also serve as the base model for the Calibrated Classifier. Calibrated Classifier modifies the predicted probabilities from the base model to ensure calibration, aligning them with the true likelihood of outcomes. Subsequently, the performance of the Calibrated Classifier is evaluated on the testing set to assess its calibration performance. This evaluation involves comparing the predicted probabilities to the observed frequencies of the positive class across different probability bins, allowing for an assessment of the model's reliability in terms of probability estimates. The optimal parameters or architecture of all the methods are presented in Table 5.

Table 5. Optimal Parameters or Architecture of all the Methods.

Method	Optimal Parameters / Architecture
Linear SVM	C = 3.745401188473625; class_weight = None; dual = True; loss = hinge; tol = 1e-05
Calibrated Classifier – Linear SVM	C = 3.745401188473625; class_weight = None; dual = True; loss = hinge; tol = 1e-05

Passive Aggressive	C = 0.1987156815341724; max_iter = 1000; tol = 1e-05
Perceptron	tol = 0.001; penalty = elasticnet; max_iter = 100; eta0 = 0.01; alpha = 0.0001
CNN	• 2 convolutional layers with 32 and 64 filters, respectively, each with batch normalization and ReLU activation in between. • Max pooling is implemented after each convolutional layer with a kernel size of 2 and a stride of 2. • Flattening layer • fully connected layer with 128 units, batch normalization, ReLU activation, and a 0.7 probability dropout sequence in order of precedence. • The final classification output is provided by another fully connected layer whose number of units is equal to the number of classes in the dataset.
DNN	• 1 fully connected layer with 128 output units. • Batch normalization applied to the output of the fully connected layer. • ReLU activation function applied to the output of batch normalization. • Dropout with a probability of 0.7 applied to the output of ReLU activation. • Another fully connected layer with 128 input units and uses the number of classes in the dataset as the output units.
LSTM	• 1 LSTM layer with hidden size of 128, with batch_first set to True • 1 fully connected layer, whose output size is equal to the number of classes in the dataset and whose input size is equal to the hidden size.
GRU	• 2 GRU layers with 128 hidden units, each followed by a dropout layer with a probability of 0.5 • 1 fully connected layer with the output size equal to the number of classes in the dataset and using the softmax activation function
MLP	max_iter = 1000; learning_rate_init = 0.01; hidden_layer_sizes = (100,); alpha = 0.01, activation = logistic

4-4- Discussion of Results

The experiment was conducted using Dataset 1 [42] and Dataset 2 [43] which underwent several preprocessing steps before being fed into machine learning or deep learning algorithms. The initial preprocessing step consisted of resizing all images in both datasets to 128 × 128 pixels. Following this, data augmentation was performed to enhance the dataset by generating 5 augmented images for each original image. Next, feature extraction was done using CNN pretrained model, specifically EfficientNetB2 and VGG16. The features extracted from both models were concatenated to create a combined feature set. The models used in the experiment included Linear SVM, Calibrated Classifier on Linear SVM, Passive Aggressive, Perceptron, CNN, DNN, LSTM, GRU and MLP. These models were trained and evaluated based on the concatenated feature set extracted from the pre-trained CNN models. The summary of experiment results is presented in Table 6.

Table 6. Summary of Experiment Results.

Method	Dataset 1 (%)				Dataset 2 (%)				Average Performance (%)			
Linear SVM	AC: 82.50	PR: 81.83	RC: 82.84	F1: 81.15	AC: 86.46	PR: 87.53	RC: 86.86	F1: 85.88	AC: 84.48	PR: 84.68	RC: 84.85	F1: 83.52
Calibrated Classifier – Linear SVM	AC: 83.75	PR: 84.98	RC: 85.62	F1: 83.86	AC: 88.54	PR: 88.03	RC: 88.85	F1: 87.84	AC: 86.15	PR: 86.51	RC: 87.24	F1: 85.85
Passive Aggressive	AC: 82.50	PR: 82.01	RC: 82.92	F1: 81.51	AC: 87.50	PR: 86.95	RC: 87.60	F1: 86.46	AC: 85.00	PR: 84.48	RC: 85.26	F1: 83.98
Perceptron	AC: 87.50	PR: 90.00	RC: 88.00	F1: 87.00	AC: 85.42	PR: 87.00	RC: 85.00	F1: 85.00	AC: 86.46	PR: 88.50	RC: 86.50	F1: 86.00
CNN	AC: 87.50	PR: 90.02	RC: 87.50	F1: 87.73	AC: 87.50	PR: 89.95	RC: 87.50	F1: 87.54	AC: 87.50	PR: 89.99	RC: 87.50	F1: 87.64
DNN	AC: 91.25	PR: 92.00	RC: 91.00	F1: 91.00	AC: 85.42	PR: 88.00	RC: 85.00	F1: 86.00	AC: 89.58	PR: 87.00	RC: 88.00	F1: 88.50
LSTM	AC: 86.25	PR: 88.03	RC: 86.25	F1: 86.25	AC: **89.58**	PR: 91.19	RC: 89.58	F1: 89.78	AC: 87.92	PR: 89.61	RC: 87.92	F1: 88.02
GRU	AC: 85.00	PR: 85.40	RC: 85.15	F1: 84.11	AC: 90.63	PR: 90.90	RC: 90.73	F1: 90.20	AC: 87.81	PR: 88.15	RC: 87.94	F1: 87.15
MLP	AC: **92.50**	PR: 93.25	RC: 93.82	F1: 92.92	AC: **89.58**	PR: 89.66	RC: 90.31	F1: 89.38	AC: **91.04**	PR: 91.46	RC: 92.07	F1: 91.15

Based on Table 6, it is clearly shown that the MLP model achieved the highest accuracy among all the models tested in both Dataset 1 and Dataset 2 with an accuracy of 92.50% and 89.58% respectively. Meanwhile, the MLP model also achieved the highest average performance among all the models with an average accuracy of 91.04%. This high accuracy can be attributed to several key factors including model complexity, utilization of non-linear activation functions and the effectiveness of parameter tuning techniques. In terms of model complexity, the MLP model is a particular kind of feedforward neural network with multiple layers between the input and output layers. In accordance with this architecture, the MLP is able to identify complicated associations in the data, which is more than it can do with simpler models like Linear SVM or Perceptron, which are unable to handle complex patterns with the same depth and complexity. Additionally, MLP can recognize and depict intricate patterns and correlations in the data because it employs non-linear activation functions in the hidden layers, such as the logistic (sigmoid) function. In addition, MLP typically have more parameters compared to simpler models like SVM or Perceptron. This increased parameterization, coupled with the careful tuning of hyperparameters such as the size of hidden layers, activation functions, and learning rate can further enhance the performance of MLP models. The parameters of the MLP model can be referred from Table 4.

The performance of the MLP model in this experiment can be attributed to its tailored adaptation to the characteristics of the datasets used. Primarily, the logistic activation function in the MLP is essential because it makes good use of the model's non-linear properties. This is especially helpful for animal classification tasks, where features frequently show non-linear separability. This enables the MLP to capture complex patterns and relationships in the data, something that linear models like SVM are less able to do. Besides, the flexibility inherent in the MLP's model architecture is also one of the key factors. The MLP model can be tailored to meet the complexity of the datasets by modifying the number of hidden layers and neurons per layer, which results in optimal performance. In this expeiment, the size of hidden layers is set to 100 after random search is performed. Last but not least, MLP have proven effective in a range of image classification tasks, which makes them a good option for recognizing endangered species in images. While CNNs are frequently employed in image classification tasks, MLP can also function well in these scenarios, particularly when paired with efficient preprocessing methods like data augmentation and resizing. This highlights how flexible and resilient MLP is when it comes to image-based classification tasks, which confirms their effectiveness when it comes to endangered animal recognition systems. The confusion matrix of the MLP model for Dataset 1 and Dataset 2 is depicted in Figure 4 and Figure 5 respectively.

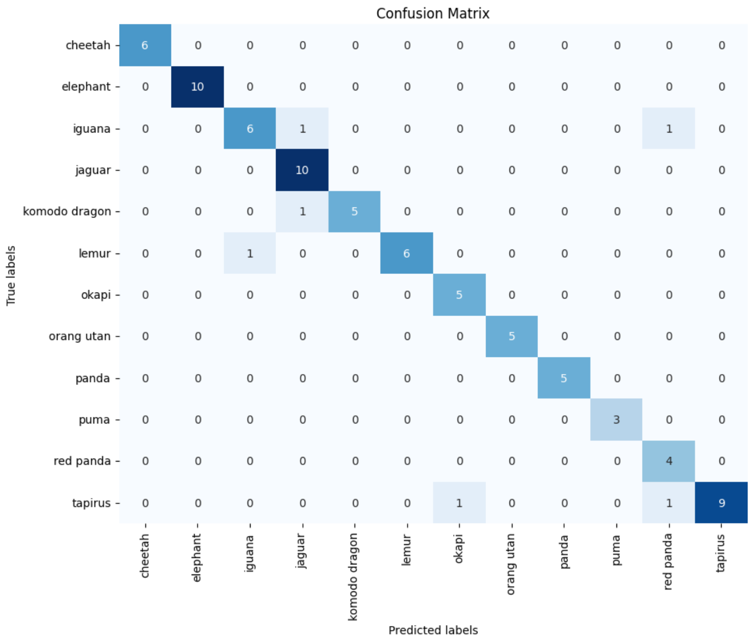

Figure 4. The confusion matrix of the MLP model for Dataset 1.

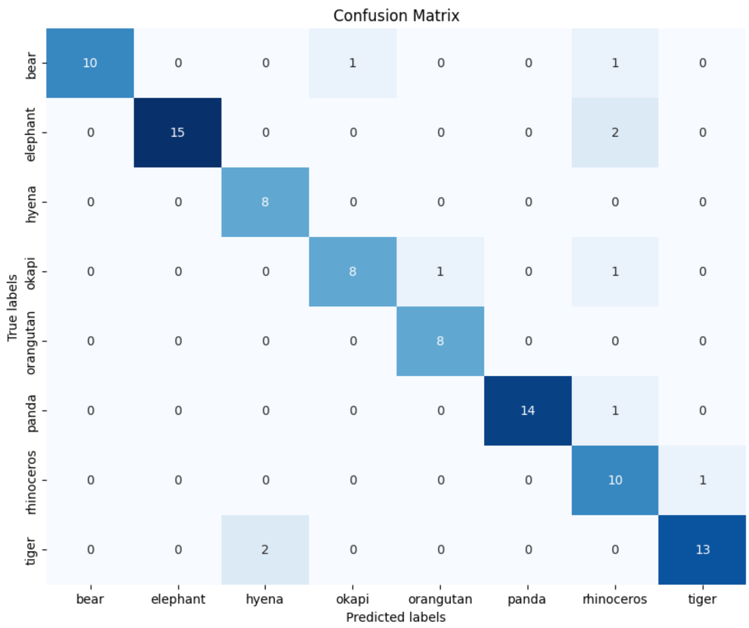

Figure 5. The confusion matrix of the MLP model for Dataset 2.

However, Linear SVM achieved the lowest average performance among all the models with an average accuracy of 84.48%. This can be explained by a number of important aspects including linear separability assumption, feature representation and dataset complexity. Linear SVM can only generate linear decision boundaries to classify data points since it assumes that the data is linearly separable by nature. It is possible that linear decision boundaries will not be enough to correctly separate classes in complex datasets with non-linear relationships, like animal classification tasks, which would result in decreased performance. Besides, unlike the MLP model, Linear SVM is unable to automatically extract features from raw data or learn hierarchical representations; instead, it is completely dependent on the feature representation that is supplied to it. This restriction can make it more difficult for Linear SVM to successfully differentiate between several classes in tasks like image classification, where features cannot be readily separable in the input space. Last but not least, linear separation might not be the best option given the intricacy of the experiment's animal classification datasets. These datasets have complex feature relationships and patterns, which makes it difficult for a linear model like SVM to categorize them correctly.

4-5- Enhancement

For enhancement being implemented in this paper, utilization form different CNN pretrained model which are EfficientNetB2 and VGG16 are adopted for the experiment. These models were chosen due to their proven effectiveness in various image classification tasks. By leveraging the distinct features extracted by each of these models, a more comprehensive feature set for classification model can be developed easily. The features extracted from both models are then concatenated to form a robust and diverse feature set. These combined features set are used to train the classification model, resulting in improved accuracy and generalization capabilities. This approach of using combined features from different CNN models proved to be a significant enhancement to this research. It illustrates the potential of feature aggregation in improving the performance of predictive models in the field of image classification.

Moreover, Lazy Predict and Random Search are also incorporated to identify the most suitable machine learning model and its parameters for training. Lazy Predict is a tool that simplifies the process of fitting and evaluating a wide range of machine learning models on a given dataset, which significantly eases the model selection process. Random Search was employed to optimize the hyperparameters of the chosen model. Unlike traditional grid search, Random Search selects random combinations of hyperparameters to train the model, which can often lead to better performance in a shorter amount of time. This method proved to be efficient and effective in determining the optimal set of hyperparameters for the model. As such, the combination of Lazy Predict and Random Search not only streamlined the model selection and training process but also contributed to the overall performance of the predictive model.

5- Conclusion

In conclusion, recent advancements in Computational Intelligence (CI) have significantly contributed to the achievement of Sustainable Development Goal (SDG) 15, specifically “Life on Land”, particularly in the field of endangered animal recognition system. A hybrid model is being proposed, which combines features extracted from pretrained models, EfficientNetB1 and VGG16, with MLP serving as the classifier. The proposed method has demonstrated promising results on two online datasets from Kaggle, excelling in terms of accuracy, precision, recall, and F1-score. Furthermore, the potential of feature aggregation is being explored to enhance the performance of the model. Looking ahead, future work will focus on investigating additional preprocessing steps, incorporating more data sources, and enhancing the model’s robustness in real-time environments. This will ensure the model’s applicability and effectiveness in practical, real-world scenarios.

6- Acknowledgements

The authors would like to extend our sincere gratitude to our lecturer, Dr. Lim Kian Meng for his invaluable advice, guidance and his consultation in answering our inquiries throughout the completion of this project. In addition, we would also like to express our gratitude to our loving parents and friends for their constant support, encouragement, and understanding throughout this project journey.

7- Contributions

As the team leader, Zhe Yu has distributed the tasks to the members. Zhe Yu also compiled everyone’s work into a final report. Zhe Yu wrote the literature review of Animal Classification System and Animal Prediction System using Machine Learning as well as completed the dataset and discussion of results for the report. Zhe Yu has also contributed in proof checking and rephrasing the final report before submission. She is also in charge of the experiment of GRU and Passive Aggressive algorithm on both datasets.

Nicole wrote the literature review of Animal Detection System Using Deep Learning, proposed method, implementation details in the final report. Nicole has also contributed in proof checking and rephrasing the final report before submission. She had also run 3 algorithms: Calibrated Classifier on Linear SVM, MLP and DNN on the experiments for 2 datasets.

Ooi Weishan wrote the introduction, literature review for Animal Classification and Poaching with Deep Learning approaches, feature extraction and experimental settings for the report. He was also in charge of running the experiments on LSTM and Perceptron in both datasets and suggested some parts in the data preprocessing steps.

Juan Kai wrote the literature review of Animal Detection System using Fuzzy System and also Animal Prediction System using Deep Learning approaches, abstract, enhancement and conclusion of the report. For the coding part, he had created a guideline or template for the team members to extract features from pretrained model and the data augmentation while training a model for CNN and Linear SVM.

8- References

[1] Pizzi, S., Rosati, F., & Venturelli, A. (2021). The determinants of business contribution to the 2030 Agenda: Introducing the SDG Reporting Score. *Business Strategy and the Environment*, *30*(1). https://doi.org/10.1002/bse.2628

[2] *The Sustainable Development Goals Report 2023: Special Edition*. (n.d.). Malaysia, Singapore and Brunei Darussalam. Retrieved January 7, 2024, from https://malaysia.un.org/en/239475-sustainable-development-goals-report-2023-special-edition

[3] Divya Meena, S., & Agilandeeswari, L. (2020). Invariant Features-Based Fuzzy Inference System for Animal Detection and Recognition Using Thermal Images. *International Journal of Fuzzy Systems*, *22*(6). https://doi.org/10.1007/s40815-020-00907-9

[4] Mohammadi, A., Park, P. Y., Mukherjee, A., & Liu, X. (2022). Developing a situation and threat assessment framework for a next generation roadside animal detection system. *IET Intelligent Transport Systems*, *16*(1). https://doi.org/10.1049/itr2.12129

[5] Kumar, R., Bhatti, S., Ruk, S. A., & Pathan, N. (2022). Identification of Endangered Animal species of Pakistan using Classical and Ensemble Approach. *2022 2nd International Conference on Computing and Machine Intelligence, ICMI 2022 - Proceedings*. https://doi.org/10.1109/ICMI55296.2022.9873801

[6] Padovese, B. T., & Padovese, L. R. (2019a). A Machine Learning Approach to the Recognition of Brazilian Atlantic Forest Parrot Species. *BioRxiv*.

[7] R, S. R. B., Nikhil, V. R., Reddy, P. V. B., S, V. P., & C, A. (2023). Endangered Bird Species Classification Using Machine Learning Techniques. *International Journal for Research in Applied Science and Engineering Technology*, *11*(5). https://doi.org/10.22214/ijraset.2023.51172

[8] Ee, K., Ab. Nasir, A. F. Bin, P. P. Abdul Majeed, A., Mohd Razman, M. A., & Ismail, N. H. (2021). The Animal Classification: An Evaluation of Different Transfer Learning Pipeline. *MEKATRONIKA*, *3*(1). https://doi.org/10.15282/mekatronika.v3i1.6680

[9] Junior, L. C. S., Padua, M. B. S., Ogusuku, L. M., Albertini, M. K., Pimentel, R., & Backes, A. R. (2020). Wild boar Classification by Using a Fusion of Texture Analysis Methods. *Proceedings - 2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images, SIBGRAPI 2020*. https://doi.org/10.1109/SIBGRAPI51738.2020.00048

[10] Vaishnavi, S., Udhaya Vani, P., Julie, J., & Joshan Athanesious, J. (2022). Prediction of Endangered Animals Using Linear Regression. *2022 1st International Conference on Computational Science and Technology, ICCST 2022 - Proceedings*. https://doi.org/10.1109/ICCST55948.2022.10040445

[11] de Oliveira Caetano, G. H., Chapple, D. G., Grenyer, R., Raz, T., Rosenblatt, J., Tingley, R., Böhm, M., Meiri, S., & Roll, U. (2022). Automated assessment reveals that the extinction risk of reptiles is widely underestimated across space and phylogeny. *PLoS Biology*, *20*(5). https://doi.org/10.1371/journal.pbio.300154

[12] Maalini, M. D., Nivetha, S., Priya, C., & Sangavi, P. (2023). DETECTION OF ANIMALS IN AGRICULTURAL LAND USING CNN ALGORTHIM. *INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)*, *03*, 1086–1090. www.ijprems.com

[13] Sreedevi, K. L., & Edison, A. (2022). Wild Animal Detection using Deep learning. *INDICON 2022 - 2022 IEEE 19th India Council International Conference*. https://doi.org/10.1109/INDICON56171.2022.10039799

[14] Altobel, M. Z., & Sah, M. (2021). Tiger Detection Using Faster R-CNN for Wildlife Conservation. *Advances in Intelligent Systems and Computing*, *1306*. https://doi.org/10.1007/978-3-030-64058-3_71

[15] Pillai, R., Gupta, R., Sharma, N., & Bansal, R. K. (2023). A Deep Learning Approach for Detection and Classification of Ten Species of Monkeys. *International Conference on Smart Systems for Applications in Electrical Sciences, ICSSES 2023*. https://doi.org/10.1109/ICSSES58299.2023.10199762

[16] Ma, D., & Yang, J. (2022). YOLO-Animal: An efficient wildlife detection network based on improved YOLOv5. *2022 International Conference on Image Processing, Computer Vision and Machine Learning, ICICML 2022*. https://doi.org/10.1109/ICICML57342.2022.10009855

[17] Teng, C. Y., Connie, T., Choo, K. Y., & Goh, M. K. O. (2022). A Visual Approach Towards Wildlife Surveillance in Malaysia. *2022 10th International Conference on Information and Communication Technology, ICoICT 2022*. https://doi.org/10.1109/ICoICT55009.2022.9914861

[18] Choe, D. G., & Kim, D. K. (2020). Deep Learning-based Image Data Processing and Archival System for Object Detection of Endangered Species. *Journal of Information and Communication Convergence Engineering*, *18*(4). https://doi.org/10.6109/jicce.2020.18.4.267

[19] Castañeda, J. A. J., De Castro, A. L., Sy, M. A. G., AlDahoul, N., Tan, M. J. T., & Karim, H. A. (2023). Development of a Detection System for Endangered Mammals in Negros Island, Philippines Using YOLOv5n. *Lecture Notes in Electrical Engineering*, *983 LNEE*. https://doi.org/10.1007/978-981-19-8406-8_35

[20] Binta Islam, S., Valles, D., Hibbitts, T. J., Ryberg, W. A., Walkup, D. K., & Forstner, M. R. J. (2023). Animal Species Recognition with Deep Convolutional Neural Networks from Ecological Camera Trap Images. *Animals*, *13*(9). https://doi.org/10.3390/ani13091526

[21] Cardoso, A. S., Bryukhova, S., Renna, F., Reino, L., Xu, C., Xiao, Z., Correia, R., Di Minin, E., Ribeiro, J., & Vaz, A. S. (2023). Detecting wildlife trafficking in images from online platforms: A test case using deep learning with pangolin images. *Biological Conservation*, *279*. https://doi.org/10.1016/j.biocon.2023.109905

[22] Rajeshwari, D., & Author, T. (2023). *Advances in Ecological Surveillance: Real-Time Wildlife Detection using MobileNet-SSD V2 CNN*. https://doi.org/10.21203/rs.3.rs-3666180/v1

[23] Sajithra Varun, S., & Nagarajan, G. (2023). DeepAID: a design of smart animal intrusion detection and classification using deep hybrid neural networks. *Soft Computing*. https://doi.org/10.1007/s00500-023-08270-1

[24] Pillai, R., Sharma, N., & Gupta, R. (2023). Classification of Endangered Species of Wild Cats through Xception Transfer Learning Model. *2023 International Conference in Advances in Power, Signal, and Information Technology, APSIT 2023*. https://doi.org/10.1109/APSIT58554.2023.10201732

[25] Tan, M., Chao, W., Cheng, J. K., Zhou, M., Ma, Y., Jiang, X., Ge, J., Yu, L., & Feng, L. (2022). Animal Detection and Classification from Camera Trap Images Using Different Mainstream Object Detection Architectures. *Animals*, *12*(15). https://doi.org/10.3390/ani12151976

[26] Shanthakumari, R., Nalini, C., Vinothkumar, S., Govindaraj, B., Dharani, S., & Chindhana, S. (2022). Image Detection and Recognition of different species of animals using Deep Learning. *2022 International Mobile and Embedded Technology Conference, MECON 2022*. https://doi.org/10.1109/MECON53876.2022.9752203

[27] Mou, C., Liang, A., Hu, C., Meng, F., Han, B., & Xu, F. (2023). Monitoring Endangered and Rare Wildlife in the Field: A Foundation Deep Learning Model Integrating Human Knowledge for Incremental Recognition with Few Data and Low Cost. *Animals*, *13*(20). https://doi.org/10.3390/ani13203168

[28] Ornek, A. H., Sen, D., & Civil, E. (2023). *SyncDreamer for 3D Reconstruction of Endangered Animal Species with NeRF and NeuS*. http://arxiv.org/abs/2312.13832

[29] Feng, J., & Li, J. (2022). An Adaptive Embedding Network with Spatial Constraints for the Use of Few-Shot Learning in Endangered-Animal Detection. *ISPRS International Journal of Geo-Information*, *11*(4). https://doi.org/10.3390/ijgi11040256

[30] Shekapure, R., Sawant, S., Kadam, S., & Chaudhari Chaudhari, S. S. (2023). Endangered Species Conservation. *International Research Journal of Engineering and Technology*. www.irjet.net

[31] Zizka, A., Andermann, T., & Silvestro, D. (2022). IUCNN – Deep learning approaches to approximate species' extinction risk. *Diversity and Distributions*, *28*(2). https://doi.org/10.1111/ddi.13450

[32] Mukadam, M., Jayaram, M., & Zhang, Y. (2020). A Representation Learning Approach to Animal Biodiversity Conservation. *COLING 2020 - 28th International Conference on Computational Linguistics, Proceedings of the Conference*. https://doi.org/10.18653/v1/2020.coling-main.26

[33] Islam, S. B., & Valles, D. (2020). Identification of Wild Species in Texas from Camera-trap Images using Deep Neural Network for Conservation Monitoring. *2020 10th Annual Computing and Communication Workshop and Conference, CCWC 2020*. https://doi.org/10.1109/CCWC47524.2020.9031190

[34] Sahil Faizal, & Sanjay Sundaresan. (2022). Wild Animal Classifier Using CNN. *International Journal of Advanced Research in Science, Communication and Technology*. https://doi.org/10.48175/ijarsct-7097.

[35] Tabak, M. A., Norouzzadeh, M. S., Wolfson, D. W., Newton, E. J., Boughton, R. K., Ivan, J. S., Odell, E. A., Newkirk, E. S., Conrey, R. Y., Stenglein, J., Iannarilli, F., Erb, J., Brook, R. K., Davis, A. J., Lewis, J., Walsh, D. P., Beasley, J. C., VerCauteren, K. C., Clune, J., & Miller, R. S. (2020). Improving the accessibility and transferability of machine learning algorithms for identification of animals in camera trap images: MLWIC2. *Ecology and Evolution*, *10*(19). https://doi.org/10.1002/ece3.6692

[36] Huang, Y. P., & Basanta, H. (2021). Recognition of Endemic Bird Species Using Deep Learning Models. *IEEE Access*, *9*. https://doi.org/10.1109/ACCESS.2021.3098532

[37] Simões, F., Bouveyron, C., & Precioso, F. (2023). DeepWILD: Wildlife Identification, Localisation and estimation on camera trap videos using Deep learning. *Ecological Informatics*, *75*. https://doi.org/10.1016/j.ecoinf.2023.102095

[38] Choe, D., Choi, E., & Kim, D. K. (2020). The Real-Time Mobile Application for Classifying of Endangered Parrot Species Using the CNN Models Based on Transfer Learning. *Mobile Information Systems*, *2020*. https://doi.org/10.1155/2020/1475164

[39] Padovese, B. T., & Padovese, L. R. (2019b). Machine Learning for Identifying an Endangered Brazilian Psittacidae Species. *Journal of Environmental Informatics Letters*, *2*(1). https://doi.org/10.3808/jeil.201900013

[40] Kuruppu, S. (2023). *AI System to Protect Endangered Animal Population and Prevent Poaching Threats using Weapon Detection*. www.ijisrt.com

[41] Brown, D., & Schormann, D. (2019). Poacher Detection and Wildlife Counting System. *22nd Southern Africa Telecommunication Networks and Applications Conference (SATNAC), August*.

[42] Menon, S. S. (2022, February 5). *Animals 151*. Kaggle. Retrieved January 13, 2024, from https://www.kaggle.com/datasets/sharansmenon/animals141

[43] Banerjee, S. (2022, July 17). *Animal Image Dataset (90 different animals)*. Kaggle. Retrieved January 13, 2024, from https://www.kaggle.com/datasets/iamsouravbanerjee/animal-image-dataset-90-different-animals