Section 1 : Topic Submission Form

This form should be submitted by the mentioned deadline.

Name:___Karan Sanjay Kajrolkar___________________________________

Student Number:__________1089046___________________________

Course:__Master in Data Science________________________

Fill your topic/s below

Project Title/Area 1: Sound Classification on Insect Audio of species Cicadas and Orthoptera Using CNN

Dataset: https://zenodo.org/record/7828439

Description:

Animal vocalizations and natural soundscapes are fascinating objects of study and contain valuable evidence about animal behaviors, populations and ecosystems. Acoustic insect recognition can add new insights that are complementary to traditional or camera-based surveillance. Many important insect groups make species-specific sounds and sound classification could be a great tool to improve and accelerate specimen identification which helps to monitor biodiversity and species distribution. Here are hundreds of thousands of insect species, making their identification an extremely challenging task. Putting into perspective, in the Netherlands alone, 40% of total biodiversity is estimated to consist of insects, whereas mammals and plants are just 5% and 8%, respectively.

Proposed Leaf (Learnable Frontend) 2021 achieved higher performance on different datasets in (https://arxiv.org/pdf/2101.08596.pdf) a similar technique used to compare with conventional spectrogram-based audio representation in (https://arxiv.org/abs/2304.12739) & proposed better result than conventional spectrogram.

Audio files going to use for research are noisy including silent, a mixture of human & environmental sound. We are using 66 species of Cicidas and Orthoptra publically available on zenodo.org. 1554 recordings from 66 species, a total length of over 24 hours and a minimum of ten files per species with files were standardized to 44.1 kHz mono WAV files ranging in length from less than one second to several minutes. Files containing long periods without insect sounds were edited into multiple smaller files with silent periods no longer than 5 seconds.

So aim of this research is to propose a CNN model to classify insects based on an audio spectrogram to achieve higher accuracy by using sound event detection to identify the actual frequency of insects with an augmentation technique, filtering noisy audio files by removing silent frequency.

Project Title/Area 2: Bird audio classification on western Mediterranean wetland Birds Species.

Dataset: Here we used publically available dataset:

- Audio files in .mp3 format https://zenodo.org/record/7505820 included files for 1-sec windows mel -spectrogram vectors of original .mp3 files.

Description:

Bioacoustics avian life monitoring is a valuable means for obtaining relevant information regarding birdlife in a specific environment where ecologists willing to apply deep learning classification techniques for bioacoustics monitoring should take into account some issues. First, properly training a deep neural network from scratch requires huge amounts of annotated audio recordings. Their performance increases logarithmically with the volume of properly annotated training data. Whereas there exist several excellent and publicly available repositories of bird sound recordings (e.g. Xenocanto or Avibase), these often suffer from label reliability and the presence of environmental noise. So on that (Joan Gómez-Gómez 2023) made a manual labeling bird audio dataset from Xeno-canto, They selected 20 of its endemic species were selected from Aiguamolls de l'Emporda` Natural Park, located in Catalonia, in northeastern Spain.

The audio file of the dataset contains 201.6 min (12,096 s) and 5795 audio excerpts from 879 original Xeno-Canto audio files. In Xeno-Canto, audios are labeled by quality from A to E, where A means that the audio quality is excellent and E means that the audio quality is poor. Hence, only files that were labeled with categories A and B from the selected species are present in the dataset. The dataset have labeled song, call, drumming, clapping etc. for each species file. All data were collected from different locations as per availability on Xeno-Canto. With original sampling frequency.

(Joan Gómez-Gómez 2023) baseline pre-trained model MobileNetV2 achieves an average F1-score less than 5% lower than ResNet50 (0.789 vs. 0.834), but for the models seem to have struggled with more or less the same species (Ciconia ciconia, Himantopus himantopus, Acrocephalus arundinaceus). Considering that these three classes are among the categories with more seconds of samples in the dataset, these might be a warning or an indicator of potential class-imbalance problems. But the combination of different spectrogram configurations and data augmentation techniques on the Western Mediterranean Wetland Birds dataset, such as noise mixing, time shifting, or mix-up help to increase the accuracy of poorly classified species. Also Pruning and Quantization techniques helps CNN to increase accuracy and reduce a loss (doi: 10.1109/AFRICON51333.2021.9570862) to propose a higher accuracy model than the baseline model.

Project Title/Area 3: Biodiversity Snake: Image classification using deep learning

Dataset:

- https://ptak.felk.cvut.cz/plants/plants/SnakeCLEF2023/SnakeCLEF2023-train-small_size.tar.gz
- https://ptak.felk.cvut.cz/plants/plants/SnakeCLEF2023/SnakeCLEF2023-val-small_size.tar.gz
- https://ptak.felk.cvut.cz/plants/plants/SnakeCLEF2023/SnakeCLEF2023-TrainMetadata-iNat.csv
- https://ptak.felk.cvut.cz/plants/plants/SnakeCLEF2023/SnakeCLEF2023-ValMetadata.csv

Description:

Developing a robust system for identifying species of snakes from photographs is an important goal in biodiversity but also for human health. The survey of R.M.M.K. Namal Rathnayaka (2021) on Ceylon krait highly venomous endemic species inhabiting the wet zone of Sri Lanka with similar looking species. Where they have taken 5 case studies for 2 for venomous and 3 for non-venomous patients. In the third and fourth patients, anti-venom was administered because admitting medical officer has mistakenly identified the offending snakes as kraits. Antivenom is recommended unnecessarily and this may be very risky to the patient because it is associated a high incidence of reactions especially for patient having severe asthmas WHO_Snake_bite_Guildlines (https://apps.who.int/iris/bitstream/handle/10665/204464/B4508.pdf) Page no 81. One of the possible reasons for the high mortality in snakebites is his type of inappropriate use of available antivenom.

The dataset used has 103,404 snake observations with 182,261 photographs belonging to 1,784 snake species observed in 214 countries. The data were gathered from the online biodiversity platforms – iNaturalist and Herpmapper. The provided dataset has a heavy long-tailed class distribution, where the most frequent species (Natrix natrix) is represented by 1,262 observations (2,079 images) and the least frequent species by just 3 observations.

So aim of the study making a model has to consider and minimize the danger to human life and the waste of antivenom if a bite from the snake in the image were treated as coming from the top-ranked prediction.

Fill in this section if a member of staff has agreed to be your supervisor:

Member of Staff: Akash Talwariya

If you have found a supervisor then you and the member of staff who agreed to supervise your project should sign below.

Karan Kajrolkar
Student Signature

Akash Talwariya
Supervisor Signature

____
Date

____
Date

Section 2: Topic Selection Research

Topic 1:

Table 1:

Title	**Link to the Paper**	**Dataset Link**
Adaptive Representations of Sound for Automatic Insect Recognition(2022)	https://www.researchgate.net/publication/365486443_Adaptive_Representations_of_Sound_for_Automatic_Insect_Recognition	https://zenodo.org/record/7828439
Classification of Complicated Urban Forest Acoustic Scenes with Deep Learning Models. Forests (2023)	https://www.mdpi.com/1999-4907/14/2/206	
(Survey) Computational bioacoustics with deep learning: a review and roadmap(2022)	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8944344/	

Cicada Species Recognition Based on Acoustic Signals. Algorithms(2022)	https://www.mdpi.com/1999-4893/15/10/358	https://github.com/tconnie/cicada
Deep Convolutional Neural Networks and Data Augmentation for Environmental Sound Classification(2017)	https://ieeexplore.ieee.org/document/7829341	https://zenodo.org/record/401395

Table 2:

Title	**Understanding of the Dataset**	**Understanding the Methodology Used**
Adaptive Representations of Sound for Automatic Insect Recognition (2022)	3 datasets used are InsectSet32 ,InsectSect48,InsectSet66 are publicaly available on zenodo.org . Consist of species of two families (Orthoptera & Cicadidae). Only files in WAV audio format with sample rates of 44.1 kHz or higher were included with only 20sec frame used.	Audio augmented with AddGaussianSNR, Masking, impulse responses (IRs) recorded in natural, AddColoredNoise. In the study, they compare conventional spectrogram-based audio representation against LEAF, a new adaptive and waveform-based frontend. Used models combinations such as Mel-spectrogram + CNN, LEAF(filter-bank and PCEN) + CNN, leafFB (training of the filterbank and temporal pooling parameters, but disable trainable PCEN) + CNN, leafPCEN(disabled training of the filterbank and temporal pooling parameters, but trainable PCEN) + CNN. the mel frontend improved from 78% to 82% and LEAF from 81% to 83%. They concluded LEAF frontend to adjust feature extraction parameters might be more relevant when there is only a limited number of audio examples. Evaluation done by Accuracy,F1-Score,recall,precision
Classification of Complicated Urban Forest Acoustic Scenes with Deep Learning Models Forests (2023)	There are seven categories of acoustic scenes for classification: human sound, insect sound(Cicidas), bird sound, bird–human sound, insect–human sound, bird–insect sound, and silence song. A dataset containing 7 acoustic scenes was constructed, with 1000 samples for each scene. The sampling rate of each sample was resampled to 22,050 Hz, the sampling bit rate is 16 bits, and the time	Convert the one-dimensional audio signal into a mel spectrogram. The number of FFT points was 1024, the frameshift was 512, and the number of mel filter groups was 128. Finally, the size of the mel spectrogram was 128x216. Audio’s augmented by Noise addition, Amplitude change, Time shifting, Frequency masking and Time

	duration is 3–5 s. Here used private datasets.	masking. Models used ResNet18, ResNet34 , DenseNet_BC_34, MobileNet_v2EfficientNet_b3. The research concluded with the DenseNet_BC_34 model had the best generalizability for new data, with an overall accuracy of 73.50%, which was somewhat attenuated compared to the overall accuracy of 93.81% achieved on the validation dataset. The reason may be that the same acoustic scene contains a variety of sound patterns, and the patterns in the test dataset have not yet been learned during the model training process and ultimately cannot be recognized by the model. The classification process will inevitably result in some misclassified samples of BirdHuman and Human, Human and BirdHuman, InsecBird and Bird. Evaluation done by Accuracy,F1-Score,recall, precision, overall accuracy
(Survey) Computational bioacoustics with deep learning: a review and roadmap (2022)	-	-
Cicada Species Recognition Based on Acoustic Signals. Algorithms(2022)	The dataset consists of a total of 43 sound recordings of three different cicada species, namely Magicicada cassinii,Magicicada septendecim, andMagicicada septendecula. All sound recordings were recorded in mono at a sampling rate of 44.1 kHz. Two types dataset used one with full length audio spectrogram other have individual syllables could be extracted from the full length audio's.	For pre-process on audio files used Butterworth filter to denoising, low pass filter to cut-off point 10 kHz, high pass filter to cut-off point 1 kHz. For segmentation used Härmä Syllable Segmentation Algorithm with improvement on finding syllables. Used CNN models with two sets of convolutional layers,relu activation, max-pooling layers and dropout layers. Then, the output was passed to the fully-connected layer and then flattened into a one-dimensional array.The array was passed to the dense layers with 64 and 32 filters, each followed by a dropout layer. Lastly, the output layer with the

		Softmax activation function was used to predict the class of the input sample. The CNN and pre-procced technique used in this paper have achive 93% on cut syllable and 100% on full length audio spectrogram score with propsed improved Harma Syllable Segmention. Evaluation done by Accuracy.
Deep Convolutional Neural Networks and Data Augmentation for Environmental Sound Classification(2017)	The dataset of UrbanSound8K is comprised of 8732 sound clips of up to 4 s in duration taken from field recordings. The clips span ten environmental sound classes: air conditioner, car horn, children playing, dog bark, drilling, engine idling, gunshot, jackhammer, siren, and street music. The dataset comes sorted into ten stratified folds, and all models were evaluated using 10-fold cross-validation, where we report the results as a box plot generated from the accuracy scores of the ten folds. For training the proposed CNN architecture we use one of the nine training folds in each split as a validation set for identifying the training epoch that yields the best model parameters when training with the remaining eight folds	The research aims to propose a deep CNN architecture for environmental sound classification and propose the use of audio data augmentation for overcoming the problem of data scarcity and explore the influence of different augmentations on the performance of the proposed CNN architecture. They have compared models with and without augmentation where augmentation is Time stretching, Pitch shifting (set 1), Pitch shifting (with a larger value than set 1), Dynamic range compression and Background noise. The proposed model CB-CNN has three convolutional layers interleaved with two pooling operations, followed by two fully connected (dense) layers applied strided max-pooling after the first two convolutional layers using a stride size equal to the pooling dimensions For training, the model optimizes cross-entropy loss via minibatch stochastic gradient descent. We use a constant learning rate of 0.01 Dropout to last 2 dense layers with a probability of 0.5. L2-regularization is applied to the weights of the last two layers with a penalty factor of 0.001. The model is trained for 50 epochs. For evaluation used proposed model in earlier research for the same dataset SKM (using the best parameterization identified), PiczakCNN (using the best performing model variant (LP)) and the newly proposed model in this research CNN model (SB-CNN). when training on the original dataset without augmentation achieved mean accuracy of 0.74, 0.73,

		and 0.73 for SKM, PiczakCNN and SB-CNN. And with an augmentation mean accuracy of 0.79. Also compared the difference in classification accuracy (the delta) when adding each augmentation set compared to using only the original training set where concluded pitch augmentations have the greatest positive impact on performance, air conditioner class is negatively affected by the DRC and BG augmentations. Given that this sound class is characterized by a continuous "hum" sound, often in the background, it makes sense that the addition of background noise that can mask the presence of this class will deteriorate the performance of the model. In future work the performance of the model could be improved further by the application of class-conditional augmentation during training one could use the validation set to identify which augmentations improve the model's classification accuracy for each class, and then selectively augment the training data. Evaluation done by Accuracy.

Topic 2

Table 1:

Title	**Link to the Paper**	**Dataset Link**
Western Mediterranean Wetland Birds dataset: A new annotated dataset for acoustic bird species classification(2023)	https://www.sciencedirect.com/ science/article/pii/S1574954123000432	https://zenodo.org/record/7505820
Multi-label Bird Species Classification Using Ensemble of Pre-trained Networks(2023)	https://ieeexplore.ieee.org/ document/10100519	
Classify Bird Species Audio by Augment Convolutional Neural Network (2022)	https://ieeexplore.ieee.org/ document/9799968	

An empirical investigation into audio pipeline approaches for classifying bird species (2021)	https://ieeexplore.ieee.org/document/9570862	

Table 2 :

Title	Understanding of the Dataset	Understanding the Methodology Used
Western Mediterranean Wetland Birds dataset: A new annotated dataset for acoustic bird species classification(2023)	He selected 20 of its endemic species were selected from Aiguamolls de l'Emporda` Natural Park, located in Catalonia, in northeastern Spain and done manually After the manual labeling process, the authors obtained a dataset of 201.6 min (12,096 s) and 5795 audio excerpts from 879 original Xeno-Canto audio files. In Xeno-Canto, audios are labeled by quality from A to E, where A means that the audio quality is excellent and E means that the audio quality is poor. Hence, the authors gathered only files that were labeled with categories A and B from the selected species. In this dataset, they labeled song, call, drumming, clapping etc for each species file. All data was collected from diff locations as per availability on Xeno-Canto with original sampling frequency.	For the spectrogram done 1-s window has at least one complete bird vocalization so a pattern can be obtained from the spectrogram. Added noise, converted to mel-spectrogram with a sampling frequency of 22,050 Hz. Used pre-trained image classification CNN models VGG16,ResNet50, MobileNetV2, EfficientNet-B0 Phase 1: all pre-trained models have already learned and tuned filters, fine-tuning is per- formed by freezing all layers in the body of the network and then training only the new fully-connected head as a warm-up phase using part of the spectrogram dataset. Phase 2: the original layers are unfrozen and another training done with a smaller learning rate is carried out to increase the overall accuracy. Used stratified 5-fold cross-validation strategy has been applied.

		In the baseline model research found good accuracy on lightweight model mobilenet 94.8% ± 2% and efficentnet 94.6% ± 2%. The models seem to have struggled with more or less the same species (Ciconia ciconia, Himantopus himantopus, Acrocephalus arundinaceus). Considering that these three classes are among the categories with more seconds of samples in the dataset, these might be a warning or an indicator of potential class-imbalance problems for the researchers using the dataset. Evaluation done by Accuracy.
Multi-label Bird Species Classification Using Ensemble of Pre-trained Networks(2023)	From Xenocanto bird sound online database that contains recordings of wild birds from all around the world. For data preprocessing we resample the bird call files to 16000Hz and 16-bit mono wave. The dataset is divided into train and test sets. The train data includes 1078 isolated bird calls of 10 different species with 1.5 se duration bird vocalization. The test set consists of 434 audio files containing multiple bird calls (Calls from 2 or 3 different species may be heard in the recording) that overlap one another.	They generated 2923 augmented mel spectrograms (time-masked, time-warped and frequency masked) using the experimental setup. bird calls transformed into mel-spectrograms using 256 mel-frequency bins (30ms frame size, and 10ms hop size) Pre-trained networks include Vgg16, IncepV3, IncepResnetV2 and Resnet50.Used Vgg16 as baseline models and ensemble with others to achieve high accuracy. Aggregated output probabilities use the Max operator to generate the ensemble predictions. Used categorical cross-entropy optimization was applied to the network with Adam. Sigmoid activation function was used to train the model over 100 epochs with 32 batches. The macro average value of the F1-score for the Mel spectrogram is 0.57 for the baseline model Vgg16.also evaluate combination of model Vgg16_Resnet50, Vgg16_IncepResnetV2_Resnet50 and Vgg16_IncepV3_IncepResnetV2_Resnet50 F1-score achieved 0.57 basline, 2-classifier 0.60,3-classifier 0.56. A combination of ensemble techniques does not perform well in the research. The proposed system beats the Vgg16 method, reporting an average F1-

		score of 0.60 for a 2-classifier model. Evaluation done by precision, recall and F1-measure.
Classify Bird Species Audio by Augment Convolutional Neural Network (2022)	On a publicly available dataset of 8000 audio examples. Bird species localization, region saving and augmentation, and feature extraction	First needing the spectrogram analytic kernel to learn what to class in bird species using librosa, and then it gets the system trained on features extracted. They have created CNNRestNet classification models that learn to discern between spectra and non-spectrogram that are used in audio to improve the accuracy of a complete end-to-end detection system. There was a loss of less than 0.0063, and the conditioning workouts accuracy is 0.9895 for the system, 0.9 as precision, and training and validation use 50 epochs in the system. A novel Technique made before giving audio to CNNResnet they have filtered out if bird voices were present they referred (doi: 10.1109/TIP.2016.2531289) which help them to classify similar bird spectrogram and filter out if no bird spectra there. Used Precision, recall, F1 score
An empirical investigation into audio pipeline approaches for classifying bird species (2021)	The main dataset contains 24000 samples of bird species from mainly North America. The second dataset includes bird species from southern Africa and the third dataset includes endangered bird species from south Africa constructed using the Xeno-Canto API and IUCN (International Union for Conservation of Nature) red list.	Parameters used to generate the Mel-Spectro are sampling rate of 44100,hop_length, fmin, fmax, n_mels, n_fft Augmentation techniques applied to the Mel-spectrograms include mixup, cutmix and specAug for CNN and pitch shift, noise injection, time shift and speed shift for DNN.No combination of augmentation applied. Used baseline model used VGG16, VGG19, MobileNetV2 and InceptionResnetV2 by Freezing the convolutional base and attaching a dataset-specific classifier. Training the attached classifier for a few epochs. Unfreezing the convolutional base and fine-tuning the whole network for a few epochs. The baseline model for CNN +transfer learning having MobileNet gives 84% accuracy but the augmentation technique does not effect increasing accuracy. Using model optimization and compression techniques like Pruning & Quantization, where

		Quantization combining 8 bit with pruning the baseline models accuracy increased by 13%, precision and recall also increased by about 5% and reduced by 96% model size which is good for deploying in edge devices. Data augmentation does not provide good enough results But the Pruning and Quantization technique helps CNN to increase accuracy and reduce loss and model size in CNN. Evaluation done by precision, recall, F1-measure and accuracy .

Topic 3

Table 1:

Title	**Link to the Paper**	**Dataset Link**
An artificial intelligence model to identify snakes from across the world: Opportunities and challenges for global health and herpetology (2022)	https://doi.org/10.1371/journal.pntd.0010647	https://lindat.mff.cuni.cz/repository/xmlui/handle/20.500.12800/1-4773
A comparative study on image-based snake identification using machine learning (2021)	https://doi.org/10.1038/s41598-021-96031-1	
(Survey) Artificial intelligence-based snakebite identification using snake images, snakebite wound images, and other modalities of information: A systematic review (2023)	https://doi.org/10.1016/j.ijmedinf.2023.105024	
EfficientNets and Vision Transformers for Snake Species Identification Using Image and Location Information(2021)	https://www.researchgate.net/publication/354312246_EfficientNets_and_Vision_Transformers_for_Snake_Species_Identification_Using_Image_and_Location_Information	

(Survey) Paediatric cases of Ceylon krait (Bungarus ceylonicus) bites and some similar looking non-venomous snakebites in Sri Lanka: Misidentification and antivenom administration (2021)	https://doi.org/10.1016/j.toxicon.2021.04.019	

Table 2:

Title	Understanding of the Dataset	Understanding the Methodology Used
An artificial intelligence model to identify snakes from across the world: Opportunities and challenges for global health and herpetology(2022)	Dataset of 386,006 training photos covering 198 venomous and 574 non-venomous snake species from 188 countries and all continents except Antarctica. We gathered photos from online biodiversity platforms (iNaturalist and HerpMapper) and a photo-sharing site (Flickr). For species with the fewest images, we further extended the dataset by scraping data from Flickr (13% of the total). Within the training set, 772 of the world's 3,921 species (±20%) had at least ten photos. The dataset has a marked long-tailed class distribution, where the most frequent species (Thamnophis sirtalis) is represented by 22,163 photos and the least frequent by just ten. For testing, they used the test set from the SnakeCLEF2021 competition with 23,673 photos. They gathered two levels of the geographical label (i.e., country and continent) because geography do a crucial role to identify species. The vast majority (77%) of photos came from the United States and Canada, with 9% from Latin America and the Caribbean, 5.7% from Europe, 4.5% from Asia, 1.8% from Africa, and 1.5% from Australia/Oceania. A small proportion of photos (ca. 1–2%), particularly from Flickr, show captive	To increase the number of samples for rare classes added incorrect species labels from Flickr which help to better performance until they overwhelm the classifier. for training augmentation used RandomResizedCrop, RandomResizedCrop, Vertical Flip, Random - ShiftScaleRotate, JpegCompression, Blur, RandomBrightnessContrast, HueSaturationValue, ImageNormalization. Original images are horizontally & vertically flipped, and the image is rotated by 180°. For modeling, the technique used ImageNet-1K pre-trained model with SGD with movement 0.9, adaptive learning rate where starting LR of 0.01 was reduced by 10% on every second epoch without validation loss reduction. loss calculated by Softmax cross entropy with mini-batch size 256. Stage 2: use both data train and validation for fine-tuning with exchanging softmax cross entropy to Focal loss which focuses on hard example + One Cycle learning rate policy used. The model macro-averaged F1 score, calculated as the mean of all species F1 scores, is 92.2% and the top-1 accuracy is 96.0%. For genus recognition, the model achieves a macro-averaged F1 score of 94.9% and a 99.0% top-1 accuracy. The macro F1 country

	snakes that are kept outside of their native range (e.g., North American Pantherophis guttatus in Europe or Australian Morelia viridis in the USA).	performance, calculated as the mean of country F1 scores, is 94.2%. Further studies need to include and more systematically compare lookalike snake species occurring at the global scale. Evaluation done by Standard-accuracy ,Macro-averaged F1 score on each classes like genre,species,location
A comparative study on image-based snake identification using machine learning (2021)	594 images of the six snake species of Lar National Park were collected, including 124 images of Caucasian pit viper, 80 of Alburzi viper, 124 of Latifi's viper, 95 of dice snake, 90 of spotted whip snake, and 81 of European cat snake. The images are of different sizes, with 24-bit RGB channels. Only those images in which at least 50% of the snake body was visible in the image were involved in the dataset collected from databases https://www.calphotos.berkeley.edu and https://www.flickr.com	All Images resized to 224x224x3. Each image was converted to a vector with a length of 150,528; afterward, the vectors were converted to a matrix with 594 rows and 150,528 columns. Partitioned to 80% for the training, and 20% for the test. For the machine learning model used, 10-fold validation set; each fold with different images in train and test, compared to other experiments, to prevent the overlapping of testing and training images in each experiment. For Neural Network used, train images were randomly rotated in a range of 0 to 45 degrees and flipped both horizontally and vertically. For this study, they compared state-of-art machine learning models using (kNN, SVM and LR) with Neural network transfer learning model (VGG and mobileNet) models set up with SGD optimizer and a learning rate equal to 0.0001, as well as a momentum equal to 0.9 As a result, a combination of LDA and SVM (kernel = 'rbf') reached a test accuracy of 84%. VGG-16, the train and test accuracy of the weighted model after one run reached 96.82 and 77.78%. MobileNetV2 accuracy obtained for the train and test of the model were 99.16 and 89.99%, 99.16 and 93.33%, 99.78 and 93.33%, 99.58 and 92.50%, and finally 100.0

		and 91.67%. The evaluation matrix used Accuracy, Precision, Recall, F1-score
(Survey) Artificial intelligence-based snakebite identification using snake images, snakebite wound images, and other modalities of information: A systematic review (2023)	-	-
EfficientNets and Vision Transformers for Snake Species Identification Using Image and Location Information (2021)	The training dataset of the SnakeCLEF 2021 and AICrowd Snake Species Identification Challenge around 5 consists of 386,006 photographs of 772 snake species. Those photographs were taken in 188 countries. The photographs originated from three different data sources, two online biodiversity platforms, namely iNaturalist(n=277,025 images;71.77%)and_Herpmapper(n=58,3 51 image s; 15.12 %) and another source containing noisy data downloaded from Flickr(n=50,630 images; 13.12 %) In addition to the photographs, metadata that provides information about the continent and country of the image location were available. Most images (n=246,482; 63.85%) were recorded in the United States of America. For 50,879 (13.18%) images, no country information was provided and 51,061 images (13.23%), had no continent information. Those images were marked with the "unknown" flag.	The preprocessing stage Object Detection was implemented, which was trained to detect single snakes in the photographs. Data augmentation was used are rotation, scaling, and noise. EfficientNets and ViTs were trained to distinguish between the snake species. Finally, optional multiplication of the models prediction probabilities and the a priori probability distribution of the snake species given the location was implemented. Filtering dataset from "Image Not found", "out-of-class images", "images contain no snakes" by identify them for exclusion from the training set, a standard ImageNet classifier with 1,000 classes and based on a ResNet50 architecture has been used. With this classifier, 6,110 out-of-class images (10.47%) have been identified as out-of-class images in the Flickr dataset. The excluded images were assigned to 384 species. The filtered dataset contained images of all 772 snake species. Data augmentation has been used to increase the training image dataset by adding slightly modified copies from existing training images to the training dataset. For EfficientNet-B4 models, the data augmentation pipeline includes random cropping of the images from a size of 430x430 pixels to 380x380 pixels, random rotation in

		the range of±40 degree, a width shift, height shift, random shearing and zooming each with a factor of 0.2 as well as a random horizontal flipping. During the data augmentation procedure, missing image positions were padded with the color of the nearest pixel neighbor. Additionally, the Lanczos interpolation was used. For the ViT models, the augmentation pipeline included random cropping from an image size of 250x250 to 224x224, and a random horizontal as well as vertical flip each with a probability of 0.5. The ensemble of efficentnet-b4 and ViT-L(large model) with added binary location to the prediction probabilities gives higher performance. On test dataset ,Macro-averaging F1-scores 78.75% an macro averaging f1-score across contries 82.88% and classification accuracy 90.42 % . Evaluation done by Standard-accuracy ,Macro-averaged F1 score.
Paediatric cases of Ceylon krait (Bungarus ceylonicus) bites and some similar looking non-venomous snakebites in Sri Lanka: Misidentification and antivenom administration(2021)	In this survey, they have taken species Ceylon krait and respective similar looking species and found out dissimilarity against a location where they originated, dorsal view, symptoms, body shapes	They have taken two paediatric cases of Ceylon krait bites and three cases of similar looking non-venomous snakebites. In the study, they research on nature and size of snakes and found dorsal view of snakes did not help in identifying snakes, but symptoms like respiratory problems after bite and as they were only found in wet locations.