SNAKE SPECIES DETECTION

LITERATURE SURVEY

Machine Intelligence

BACHELOR OF TECHNOLOGY- V Sem CSE

Department of Computer Science & Engineering

SUBMITTED BY

Batch No: 12

STUDENT NAME	SRN
Amisha Mathew	PES2UG20CS038
Anurag G	PES2UG20CS057
Anushka Gupta	PES2UG20CS060

PES UNIVERSITY

(Established under Karnataka Act No. 16 of 2013)

100 Feet Ring Road, BSK III Stage, Bengaluru-560085

PAPER 1: Smart Snake Identification using Video Processing

IEEE paper: https://ieeexplore.ieee.org/document/9707360

AUTHORS:
P.D.R. Deshan, D.V.H. Pabasara, N.A. Yapa, D.S.R.C.V. Perera, Dilani Lunugalage, Janaka L. Wijekoon

INTRODUCTION:
This paper proposes an approach to help people in identifying snakes in panic situations using video processing, and then alert the nearest rescuer teams.
The implemented system comprises a mobile application with features including offline real-time snake identification, online real-time snake identification using video processing, manual snake detection, and alert nearest rescuers. The obtained results indicate that this application has an offline snake identification accuracy of 75%-80% and an online snake identification accuracy of 90%- 95%. The video processing method was considered as people found it difficult to capture photos in a panic situation. Besides, the margin for error when working with images is considerably high.

METHODOLOGY AND RESULTS:

A. Identification of snake species in the offline mode:

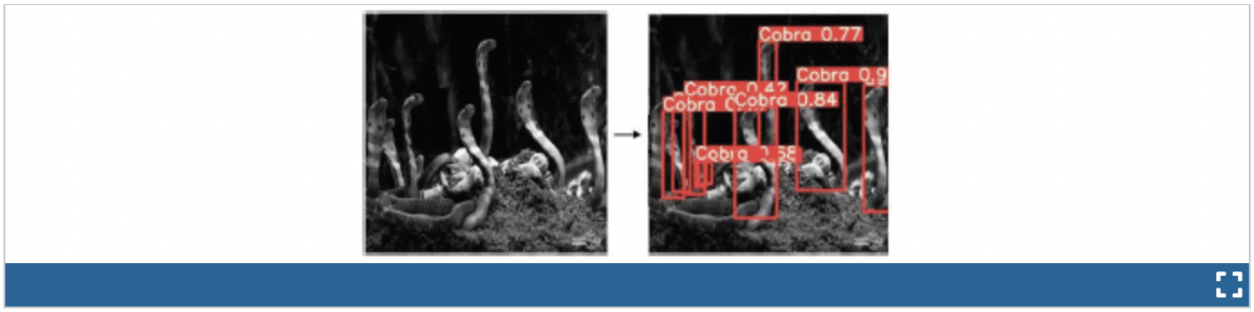
The user uploads an image to the mobile application in this to identify the snake and its venomous level. The model was trained using the You Look Only Once Algorithm (YOLO) version 5 for this process. The trained YOLOv5x model is converted to a TensorFlow Lite model which is quantized for Float16 after conversion to reduce the size of the model.

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

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

$$mAP = \frac{\sum_{\mathrm{i}=0}^{\mathrm{N}-1} \int_0^1 \mathrm{p(R)dR}}{\mathrm{N}}$$

Using the above equation, Mean Average Precision (mAP) is a good metric to evaluate object detection models and mAP value of this model was 0.89, which was near to 1 and can be considered an accurate model.



This method identified almost all the cobras.

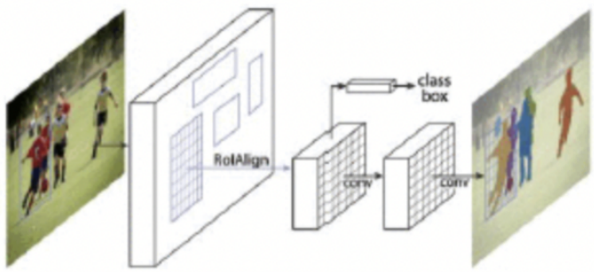
B. Video encoding and decoding:

A video is uploaded which is limited to 15 seconds to the central server. The video is encoded using FFMPEG library, using H.256 codec with libx265. This is done as a lossy compression to make the process fast. In the decoding process image frames are extracted from the video using the OpenCV library. One frame from each second is extracted.

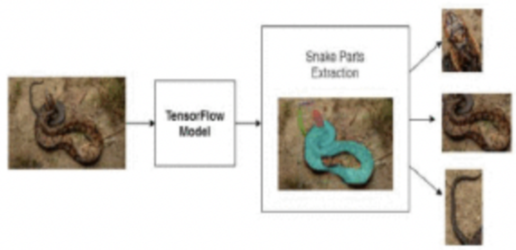
The model used in the snake parts segmentation process which is trained using Mask R-CNN represents mAP value of 0.95. which is near to 1. Therefore, can be considered as an accurate model.

C. Identification of snake species using head, tail and body pattern in server side:

This is used for higher accuracy of detections. The process use TensorFlow models which were trained based on Mask R-CNN.



Mask R-CNN model architecture



D. Manual identifications with Facebook Graph API and NLP:
The “POST” type API and the “GET” type Facebook graph API is called by the central server to read all the comments given to the Facebook post. The comment set will be analyzed using NLP to identify the snake specie and the output is generated.

E. Sending sms notification to nearby rescuers:
A crowdsourcing platform was used to maintain the rescuers' information, and Mongo DB database was used to record those details. Twilio SMS gateway API was used to send the SMS notification to the snake rescuers. Reverse Geocoding API has been used to find the location of the nearest rescuers and an sms notification is sent.

PAPER 2: Snake species identification by using natural language processing

ResearchGate paper:
https://www.researchgate.net/publication/332140002_Snake_species_identification_by_using_natural_language_processing

AUTHORS:
Nur Liyana Izzati Rusli1, Amiza Amir2, Nik Adilah Hanin Zahri3, R. Badlishah Ahmad

INTRODUCTION:
The descriptions were presented in unstructured text, and the NLP processing involves pre-processing, feature extraction and classification. Four machine learning algorithms (naïve Bayes, k-Nearest Neighbour, Support Vector Machine, and Decision Trees J48) were used during training and classification. Results show that J48 algorithm obtained the highest classification accuracy of 71.6% correct prediction for the NLP-Snake data set with high precision and recall.

METHODOLOGY:

1. Raw data collection

The respondents were asked a few questions in a questionnaire to describe the snake image that they had seen based on their perception and opinion of the snake explaining the eight physical characteristics of the snake.

2. Text pre-processing

This includes data tokenization, stemming, symbols and stop-word elimination.

3. Feature extraction using TF-IDF

High- weighting keywords are extracted and term frequency(TP) is initialized to each word.

$$Normalized\ TF = \frac{No.of\ term\ that\ occurred\ in\ the\ text}{Total\ no\ of\ word\ in\ the\ text}$$

$$IDF = \log\left(\frac{Total\ number\ of\ texts}{No.of\ text\ in\ which\ selected\ term\ is\ appeared}\right)$$

$$TF\_IDF = Normalized\ TF * IDF$$

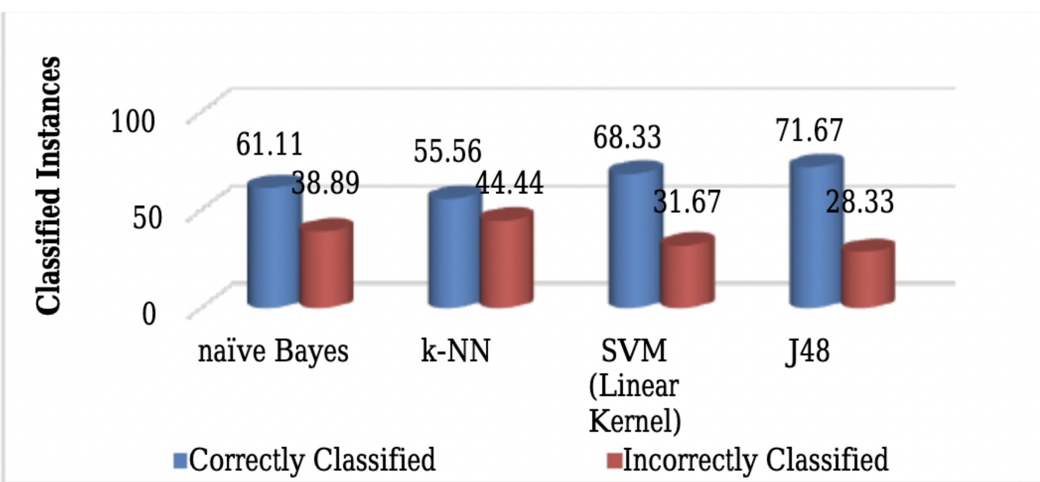
4. Training and classification

Four machine- learning algorithms- naïve Bayes, Support Vector Machine (SVM), k-Nearest Neighbours (k-NN), and decision tree. 10-fold stratified cross validation was applied as the sample was limited.

RESULTS AND ANALYSIS:

The raw dataset in text form was imported into an Attribute-Relation File Format (ARFF) file. During preprocessing, Weka package was used to convert words into the vector. The feature extraction tasks result in a reduction of the dimensionality of attributes to 30%.

Classification accuracy:

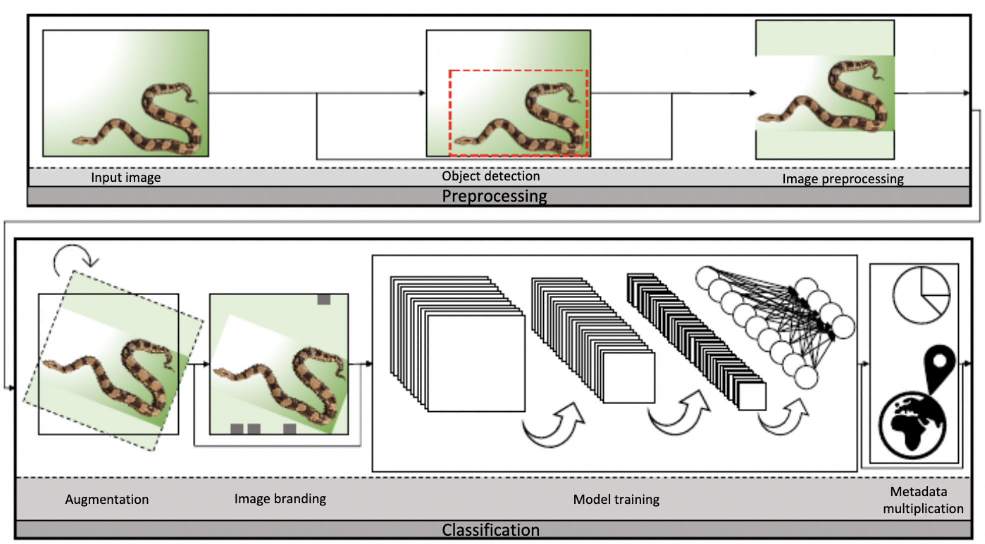


PAPER 3: Combination of image and location information for snake species identification using object detection and EfficientNets

AUTHORS:
Louise Bloch, Adrian Boketta, Christopher Keibel, Eric Mense, Alex Michailutschenko, Obioma Pelka , Johannes R¨uckert, Leon Willemeit, and Christoph M. Friedrich

INTRODUCTION:
The implemented machine learning workflow uses Mask Region-based Convolutional Neural Network (Mask R-CNN) for object detection, various image pre-processing steps, EfficientNets for classification as well as different methods to fuse image and location information. The best model achieved a macro-averaging F1-score of 0.594.

METHODOLOGY:
Python, Keras and Tensorflow models are used in this. Image pre-processing was done and the images were augmented, and fed into the deep learning training network. Finally, multiplication of the prediction probabilities and the apriori probability distribution of the snake species occuring at the corresponding location has been implemented.



1. Object Detection

This is implemented using the Mask R-CNN procedure. Mask R-CNN performs instance segmentation as it extracts a bounding box, a class label and a pixel-wise segmentation mask for each object detected in an image. This is done in two stages. The first stage uses a backbone CNN which extracts a feature map from the original image. A Region Proposal Network (RPN) is used to identify candidate object regions. The second stage consists of a Region of Interest (ROI) align network which extracts multiple possible ROI sections. Additionally, a CNN-based mask classifier is employed for pixel-wise segmentation.

2. Image pre-processing

The methods used here are resizing, scaling and filling the boundaries.

3. Data augmentation

This has been used to expand the training images and avoid overfitting. In each epoch of the training process, the images were randomly transformed.

4. Image classification

EfficientNets:

$$w_1(c) = \frac{\max F(c)}{F(c)}$$

$$w_2(c) = 1 - \frac{1}{\sqrt{\frac{\max F(c)}{F(c)} + 0.5}}$$

Polyak averaging:

$$W_{polyak}(i) = \exp\left(\frac{-i}{2}\right)$$

5. Addition of location information and image branding

$$b_w = \left\lfloor \frac{d}{8} \right\rfloor - 4$$