Journal of Theoretical and Applied Information Technology
15th February 2024. Vol.102. No 3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

PROVISIONING AN UNCERTAINTY MODEL FOR ELEPHANT MOVEMENT PREDICTION USING LEARNING APPROACHES

R. VASANTH[1], A. PANDIAN[2,*]

[1,2]Department of Computing Technologies,

[1,2]School of Computing, College of Engineering and Technology,

SRM Institute of Science & Technology,

Kattankulathur, Chengalpattu District - 603 203

Tamil Nadu, India

Mail id: vasanth1.srmit@gmail.com, vr1116@srmist.edu.in, pandiana@srmist.edu.in

ABSTRACT

This work concentrates on resolving Static/Dynamic Body movement estimation and rigid body orientation of animals. Algorithm has to be modeled with a complementary structural model that exploits measurements of the magnetometer, accelerometer, and gyroscope. Usually, attitude information is essential to evaluate animals' movement estimation to compute active sensor nodes to gather information. The risk factor is measured with Weighted Uncertainty Priority (WUP), an IVF framework. WUP considers various relative weights of complex risk factors by examining their degree of ambiguity/uncertainty. Uncertainty measure is an evidence theory used to generate exponential weight of every risk factor. This ambiguity measure shows a subjective assessment of internal coordination between sensors and animal movements. The anticipated algorithm reduces network overhead and effectually classifies samples with the pre-trained Yolov5 Network model to estimate the arrival of animals and to determine the correlation between normal and uncertain data. Feature-aware pattern modeling is done to schedule elephant movement versus risk factors. The theoretical analysis has to validate that the anticipated model outperforms other prevailing models. Simulation has been carried out in a MATLAB 2020a environment where detection accuracy is estimated as 99.5%.

Keywords: *A Priority-Based Weighted Scheduling, Weighted Uncertainty Priority, Non-Linear Filter, Uncertainty, Kalman Filter, Static/Dynamic Body Movement Estimation, Rigid Body Orientation*

1. INTRODUCTION

To reduce human-elephant conflict (HEC), monitoring elephants is essential in regions near forest boundaries. It is challenging to monitor elephants because of the way they move. Controlling the loss of elephants and human lives can be accomplished by understanding elephants' activity and behavior patterns [1]. Elephant tracking may be done safely, accurately, and economically with the help of wireless boundary-detecting devices. Camera sensing, among other methods, offers visual evidence for locating the object [2]. It is a well-known technique now in use for getting information from wildlife. Camera is an important tool for measuring elephants' movement and its surroundings. Camera's output is inherently flawed in so various ways, including lightning conditions, metrics of distance (visible partially, close, and too far), weather-related events brought on by rain and wind, shadows, animals other than elephants (the content only focuses on elephants), same, numerous images of the same animals in various positions [3]. Despite the possibility of taking millions of photos, the time-consuming process requires human interaction to extract the target's knowledge [4]. The answer is a system for automatically identifying elephants in raw photos created without human input. A camera, artificial intelligence, and developing machine vision technologies can all be used to complete image recognition automatically [5]. Deep Learning (DL), a deep neural network (DNN) based

subfield of machine learning, is aided by this. In terms of DNN structures, constitutional neural networks (CNNs) is the most well-known example. A subclass of feed-forward deep artificial neural networks is CNNs, and it has shown success in visual image analysis [6]. Generally, convolution technique extracts various aspects of the input.

Given the number of pixels that might be used to create a picture, it is crucial to keep image recognition software efficient by processing just the pixels that contain useful information [7]. Most pixels contain redundant data, including backdrop landscape, which is costly in time, processing and storage. Thus, useful information must be extracted to deploy such resources more effectively, such as data about wild creatures, vehicles, or people. In the binary classification task known as "background subtraction (BS)," in an image sequence, each pixel is assigned a tag, both in the foreground and the background image [8]. This study combines BS and DL, two machine learning techniques, to produce computer vision advancements that are very remarkable [9]. With the help of several hidden layers, DNNs map a receptive field to the output layer. Boundary-based algorithm combines DL framework with range of capabilities, is the paper's novel contribution, (i Animal detection in photos and (ii) species classification are included. Elephants are the main subject of this essay. By detecting the target with the greatest accuracy possible, BSDL would reduce HEC [10].

For this proposed research project, HEC is the realistic application used. The plan is to incorporate a system that will notify authorities when elephants are spotted, allowing them to evict the pachyderm and assist the public in protecting their lives and property who reside close to a forest's boundaries by chasing it back into the forest [11] – [12]. Applications in the real world also include environmental monitoring and border enforcement systems. The literature indicates that DL has received much focus and has been employed successfully in several fields, notably image processing, bioinformatics, computer security, and gaming [13]. A unique finger vein detection DL technique is shown as another way that DL is utilized to prevent biometric authentication from leaking unprotected data. DL also implements a permission usage-based malware detection system [14]. The best we can tell is the initial effort to apply DL's advantages to a realistic problem—identifying elephants near the forest's border [15]. The proposed work's motivations include automatic species detection from camera trap data and elephant image detection from other pictures of various animals. The literature claims cameras were crucial for detecting elephants by employing various database comparisons and image processing techniques. False alarm inclusion, however, must be avoided. The proposed study aims to reduce HEC through effective camera image detection that uses the BS and DL advancements in technical contributions to reduce false alarms. It is crucial to emphasize that the success of deep learning models in predicting elephant movements is contingent upon the accessibility and quality of data. Furthermore, careful attention must be given to ethical and privacy considerations associated with wildlife tracking and conservation initiatives. Collaborative efforts involving data scientists, conservationists, and domain experts are essential to create precise and ethically sound models. The major research contributions are listed below:

- This work proposes a novel approach for elephant movement detection relies on movement confidence mapping and uncertainty analysis;
- This work employs WUP model to quantify movement uncertainty for detection and further validation;
- The confidence mapping is experimented and enhances the reliability using various performance metrics as the uncertainty analysis is not executed in other fields yet. The model attains better prediction reliability.

The work is structured as: section 2 provides a comprehensive review on diverse prevailing approaches; section 3 elaborates uncertainty-based elephant movement prediction model and outcomes are elaborated in section 4. Summary is provided in section 5.

2. RELATED WORKS

Forest lands were widely utilized for human habitation due to the constantly growing human population, industrial growth, and agricultural needs [16]. Elephant death rates are steadily rising due to the extensive stretches of linear facilities like railroads and highways inside forest regions. In our planet's ecological system, elephants play a

significant role. Since they move slowly, they must travel far to satisfy their biological requirements [17]. There are still difficulties in finding these pachyderms in the wild. In many of these cases, animal crossings have made the areas vulnerable to accidents at locations with forest scenery used to develop linear systems. Accidents involving elephants on railroads are one of the serious problems that India is currently experiencing and a big source of concern [18]. Elephants commonly encounter these threats. Therefore, conserving them is essential. Therefore, the research must address the issue of detecting elephant movement within the forest area [19].

Different sensing modalities, including seismic, radar, imaging-based, and acoustic, can detect the movement of ground targets, such as people, vehicles, and animals. There are two types of these detecting methods: passive and active. Seismic sensors can detect the ground waves produced by an elephant's movement [20]. Ground-based sensor recognition and target detection have extensively used seismic sensing, a passive detection approach. Seismic sensing has the advantage of being resistant to environmental disturbances like heat, rain, and wind [21]. It can also operate in both day and nighttime circumstances. Nevertheless, seismic sensor-based systems are viewed as answer owing to capabilities for stealth detection. Other passive techniques, including thermal and acoustic imaging, are also being utilized to locate elephants [22]. In contrast to acoustic modality, seismic modality is similarly less susceptible to the Doppler effect (DE) and offers omnidirectional detection capacity for moving objects at considerable distances [23] – [24].

Several researchers have studied the vocalization and behavioral traits of elephants. Initially, Rodwell et al. tested acoustic and seismic sensors on elephants. A prominent frequency in the 20 Hz band was found when the elephant vocalization's seismic signal was analyzed. Within 10 meters of the sensor, there is a 25 Hz frequency [25] utilizing a geophone sensor to examine the features of seismic waves produced by elephant vocalization. Using a 10- 40 Hz corner frequency and 23rd-order Butterworth bandpass filter, Wood et al. used spectral and temporal patterns to determine species using the seismic features of elephant movement [26]. Elephant detection systems have been created utilizing frequency and time-based features that were retrieved from the elephant locomotion's seismic signatures by Sugumar et al. after studying the elephant's migration pattern movements. According to Mortimer et al., elephant stomping causes seismic waves, and they investigated the effects of background seismic noise and terrain type on the transmission of the signals. The vocalization and maximal elephant movement propagation range have been calculated using a simulation of the movement of waves with a frequency of 4.5 - 25 Hz. For artificial waveforms, the maximum propagation distance was also determined using the STA/LTA detection technique [27]. The accurate detection of an elephant was already put forth utilizing a hazy cognitive map and an acoustic, seismic, and image-based multi-modal wireless integrated sensor network.

The material already written has mostly concentrated on the conventional evaluation of elephant identification based on time. Nevertheless, several innovative frequency, time-frequency and time-domain are presented for identifying moving targets. The signal energy is often subjected to a threshold in a time-based technique to detect the target. Nevertheless, the frequency-based method detected the signal spectrum's energy. To detect moving objects, several time-frequency representations like Fourier transform (SFT), wavelet transform (WT) and Wigner-Ville distribution (WVD) utilized in time-frequency analysis, have been employed [28].

Despite numerous sporadic research initiatives, most experiments are conducted with a single elephant. Only a few researches have been documented for characterizing time-frequency of seismic vibrations of elephant. In order to raise early alerts, these applications also need automatic signal-detecting techniques. Additionally, the body of existing work has motivated the adoption of quantitative analysis based on the distance to detect and recognize elephants; for automatic elephant detection, which necessitates in-depth knowledge of the elephants' locomotion frequency, selecting a suitable filter band is essential [29]. Additionally, a discussion has been had on the performance comparison of several detection techniques for characterizing the locomotion of elephants with differences in detecting group sizes and distances. In order to emphasize detection in a forest surrounding: a study of its viability, this paper will use the

elephant locomotion's seismic properties [30]. The statistical performance indicators of precision, accuracy, false-positive rate, F1-score, and true positive rate are used to evaluate the detection performance.

3. METHODOLOGY

The uncertainty analysis module (UAM), prediction module (PM) and evaluation module are all part of the general layout. Fig 1 depicts the intended assessment process for the identification of elephant movement. An online available is fed into an uncertainty model, which then uses the prediction model to compute the prediction uncertainty and sends it to the PM. The calibration findings are then forwarded to the EM for the algorithm's final evaluation. The three primary modules are shown below and Section 3 provides more details on the prediction model.

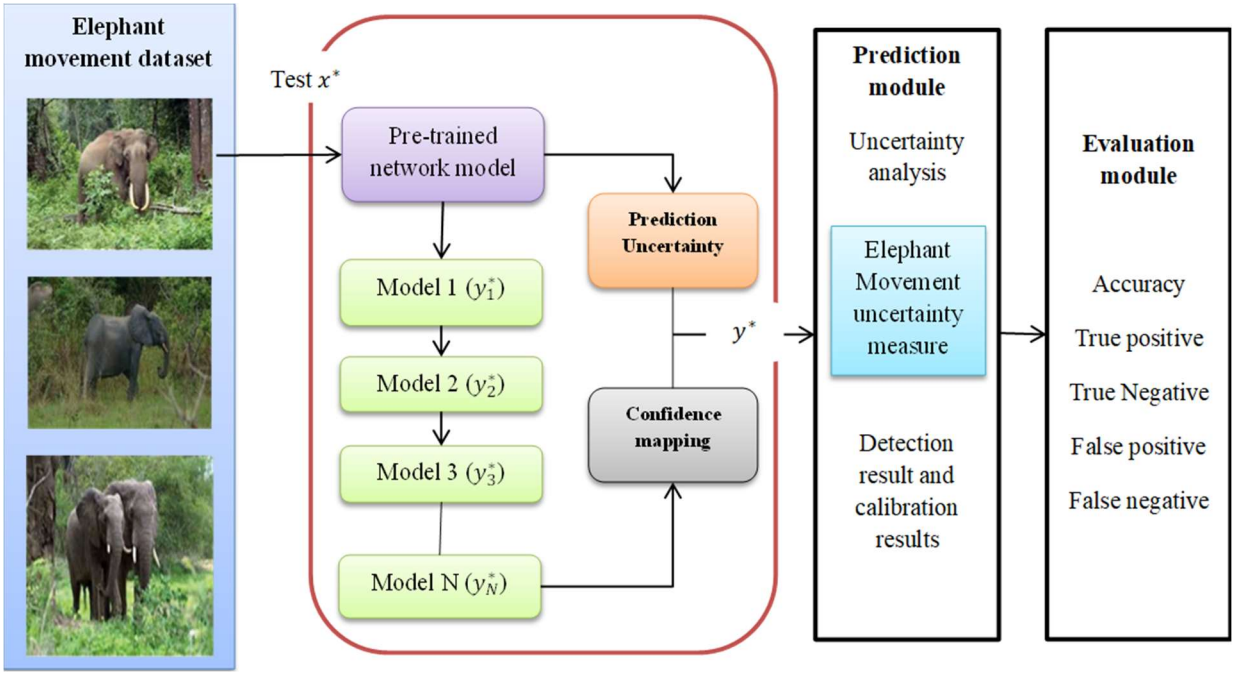

Figure 1: Prediction model

3.1. Uncertainty prediction

YOLOv5 model is tested N times via random dropout of the uncertain elephant movement data to modify the weights. As a result, using the same test data x^*, multiple detected outcomes $y^*_{(N)}$ are achieved. Identical objects' bounding boxes (elephant images) are compared using the IoU comparator shown below in various sampled detection results. IoU comparator gives each object in its cluster's sampled initial detection result. A successive cluster will only receive a bounding box if its IoU is higher than the designated threshold's IoU. Epistemic uncertainty can then be represented as necessary using the variance obtained from the boundary box cluster.

The uncertainty is defined as the mean square deviation between detected outcomes obtained using the initial variant $\{y'_V\}$ and ground truth $\{y'_V\}$ of validation set $\{x'_V\}$. The addition of the two uncertainties will yield the model's prediction uncertainty. A few calibration variables are designed in the CM, and the data from the UAM is utilized for the model's confidence level, which must be calibrated. The evaluation indices at all levels are computed following the acquisition of the calibrated result. When compared to the ground truth and associated uncertainties, all identified boxes in the assessment module are categorized as false positive (FP), false negative (FN), true negative

(TN), or true positive (TP). Finally, algorithm's performance in scenarios is assessed using AP, precision, and recall metrics.

3.2. Prediction model

CNN with pre-trained YOLOv5 contains ambiguities that are challenging for humans to understand. Bayesian inference is now the most widely used technique in DL model uncertainty analysis. The Bayes theorem is expressed as:

$$p(y^*|x^*,X,Y) = \int p(y^*|x^*,f)p(X,Y)df \quad (1)$$

Here, x^* is an input data point, y^* refers associated output and f denotes function. X and Y are the training sets of data and ground truth label respectively. The expression $p(y^*|x^*,X,Y)$ expresses the uncertainty in the model output and explains output distribution y of an unidentified model trained from X and Y. The given model distribution f output y^* is described as $p(y^*|x^*,f)$, which can be directly estimated. The expression $p(f|X,Y)$ denotes the posterior distribution of model f, which is the distribution of model f as trained by provided training sets X and Y. Limiting weight parameter W helps simplify the model f notion, which is extremely abstract.

$$(y^*|x^*,X,Y) = \int p(y^*|x^*,f)p(X,Y)df \quad (2)$$

Consequently, using the method mentioned earlier, the problem with uncertainty is changed by calculating model weight's posterior distribution. The model uncertainty is then measured using the proposed method:

$$W = M * diagonal\left(\left[z_{i,j}\right]^{K_i}_{j=1}\right), z_{i,j}B(1,p_i) \quad (3)$$

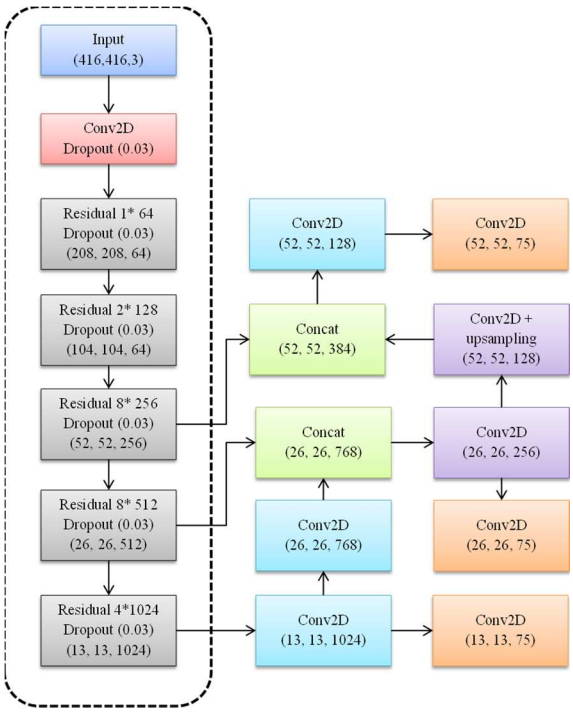

Figure 2: Yolov5 model for detection

The foundation of the proposed method is to approximate the posterior distribution $p(W|X,Y)$ using multiple Bernoulli distributions $q(W)$, where M specifies weight matrix of original model and $z_{i,j}$ describes model's j^{th} input neuron in layer $'i'$, which probabilistically follows the Bernoulli distribution, p_i. The matching neuron is inactivated when $z_{i,j}$ has a zero value. Weight matrix is created after this processing. There will be various outcomes for a given input. The inclusion of dropouts would alter the expected results of the model. As an illustration, neuron outputs 1 by default outputs 0 with probability $1-p_i$ and 1 with probability p_i. To retain the initial expectation, the output of every neuron should be split by p_i; utilizing the Monte Carlo approach to produce several sample models is another essential component of the method. After sampling the reference model weight N times, the output results are Ny^*. The variability of the data can be used to estimate the model's uncertainty:

$$Var\left(y^*_{(n)}\right) = \frac{1}{N}\sum_{n=1}^{N}\left(y^*_{(n)} - \bar{y}^*\right)^T\left(y^*_{(n)} - \bar{y}^*\right) \quad (4)$$

The model utilized in this study is YOLOv5 pre-trained, and the number of model samples is 20. The input is rearranged into 416×416 via the YOLOv5 model, which uses the YOLOv3

method. Since batch normalization has taken the position of the dropout layer in YOLOv3, the network design of YOLOv5 has been manually adjusted depending on the MCD approach, as illustrated in Fig. 2. YOLOv5 was also re-implemented to assess model uncertainty using the proposed technique. They employed the deterministic feature tensor to increase efficiency. It only added dropout layers toward other network end whereas we simulated entire model uncertainty by including dropout levels after each convolution module.

The full connection layer's dropout rate of 0.5 has reportedly been utilized to build more random network architectures and address the over-fitting issue while utilizing dropout as the regularization method, according to [20]. The last prediction network layer, $the\ 1 \times 1$ convolution layer, contains parameters than earlier convolution layers and same effects are encountered in full connection layer. Because of this, the 1×1 convolution layer's dropout rate is set to $1 - p_i = 0.5$. The convolutional shared-filter architecture drastically reduces the number of parameters for the other convolution layers, but this results in a severe loss of features due to the high dropout rate. After testing on a few pictures, the dropout rate of the remaining convolution layers has been set to 0.03. These picked rates are also hyper-parameters that can be adjusted further.

3.3. Confidence mapping

The proposed WUP is an easy non-parametric calibration technique. It separates all un-calibrated potential confidence values c_i into mutually exclusive m elephants with a calibration score denoted as c_m specified for each prediction. The calibration result will be denoted as c_n if the output is in the n^{th} prediction. The uncertainty analysis results should be used to calibrate the confidence. The WUP can extract the variance and mean of each data point in diverse sampling models. Therefore, the initial predicted results in k times the mean and SD, i.e., original model's forecast results' distribution concerning the mean value may be determined. Consequently, the value of k can be utilized to divide the elephant herds.

The calibration factor increases, and the model output becomes more stable for original test outcome with the mean sample. If the output is consistent and confident, the upper confidence limit is set as 1 once the model has been calibrated. The assumptions before and after the calibration are crucial to reducing further consequences. During uncertainty analysis, visual tasks may be regarded as obeying an a priori normal distribution. As shown in Fig 3, the calibration factors are created using the normal probability distribution. During the experiments, it was found that the sampled models sometimes failed to recognize the object that the original model had picked up. The algorithm's uncertainty is determined during design by computing $M\ (M \leq N)$ discovered sampling outcomes of certain object. Some samples are employed in specific setting, yet an elephant will only be recognized twice, which suggests substantial model uncertainty. The statistical value understates its uncertainty because both results are extremely near. The calibration, therefore, takes the M factor. Lower output confidence must be expected the smaller M is. The object (elephant) that the sampling models may also identify simultaneously that the original model did not detect.

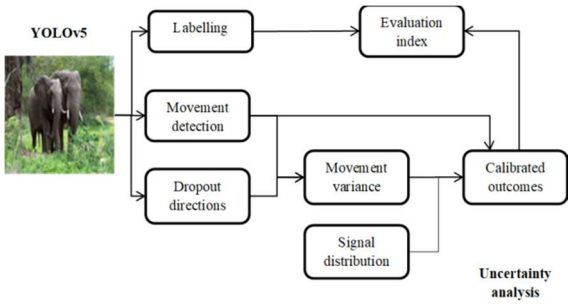

Figure 3: Confidence mapping

These scenarios are not considered because a range of uncertainty is introduced to the original model output. Mean value has been used as the model results where WUP technique is incorporated to create Bayesian model. Although this kind of technology cannot accomplish one-step object detection, semantic segmentation can. As opposed to the latter, a regression issue with difficulty during the bounding box alignment discovered by various sampling models, the former fundamentally involves the classification of pixels. Existing researchers investigated various techniques for grouping bounding boxes relies on semantic affinity and spatial discovered IoU affinity measure in combination with the sequential method. Various MATLAB 2020a

functions are built to implement the confidence. Initial prediction, final results, and 20 sampling detection findings upon calibration are achieved after inputting the ground truth. In Fig 2, the data flow is displayed. The formula for the confidence calibration is:

$$y_c^* = \frac{M}{N} * k * y^* \quad (5)$$

3.4. Elephant position detection

Although it was indicated in Section 4, the MAP assessment system, a benchmark method to evaluate object detection algorithms, still needs improvement. While conducting some optimization taking into account the peculiarities of the object detection problem, this research employs an overall evaluation framework and the mAP evaluation system. One benefit of calibrating the confidence is that the evaluation results' trustworthiness is increased. Another method involves extending the classification task's confusion matrix to identify FN, TP, FP, and TN.

The confusion matrix is one of the most fundamental and simple tools for evaluating a classification model. The number of samples the theory predicts and the number of samples the ground truth label has as the column and the row are combined to form a matrix. The numbers on the diagonal then represent the samples with predictions that agree with the actual outcomes; better forecasts are made when the numbers are larger. Every image in classifying images has to have a unique outcome, and the result may be accurate or incorrect. The object detection task does, however, also include 2 extra scenarios, the missing prediction and repetition prediction scenarios. According to the repetition prediction, at least two bounding boxes must be constructed for bounding boxes, or each object must be generated for things that do not exist. The matching bounding box for an actual object is not formed because of the missing prediction. In order to take these scenarios into account, the confusion matrix is expanded by adding an extra column and an extra row to track the missing and recurring assumptions accordingly.

To reflect the categorization effect of the technique, the original confusion matrix can still be found in the upper left corner of the expanded confusion matrix. Additionally, it may track instances of repetition and absence and assess the model's overall effectiveness. The sub-matrix is a diagonal matrix composed of the initial 4 columns and rows, also demonstrating that the model excels at classifying and will not wrongly classify a cyclist as a car. Reducing missed predictions is the hardest obstacle for object detection. Fig 3 illustrates how the approach to get indications like TN is expanded simultaneously. The elephant data subset has no incorrect detections, proving that the model's classification skill is strong. Nevertheless, repetitive forecasts and missing predictions are represented in some samples. There are five missing objects for the pedestrian category, which results in five false negatives. As a result, the FN should include 5.

4. NUMERICAL RESULTS

A thorough explanation of successive test experiments conducted. In test trials 1 through 5, various inputs were fed to various trained sets of data of elephants. In experiments 6-10, a set of data of trained elephants using the same input was used. A test experiment was performed on the trained elephant dataset using the basic elephant data from an untrained dataset. Depending on the architecture and varied accuracy limits, the WUP prediction was made (c_i for cropped and r_i for raw images). A test Experiment was carried out on the exactly trained dataset using a cropped image from the data input.

WUP prediction attained various levels of accuracy. The time ranges and accuracy when raw image data from the exact trained E1 data was utilized as feed in Test Experiment 6 were acceptable compared to the data entered from other data sets. By providing a clipped image from the trained E1 set of data as feed, the 6b test experiment was carried out using a similar E1 set of data. Fig 4a to Fd illustrates how the accuracy ranges for the WUP prediction are the same while drastically reducing the required time. A similar approach was used for other test experiments. In Experiment 7, the E2 set of data was used, followed by the E3 set in Experiments 8 and the E4 set in E9 and 10. When examining the test experiment predictions made by the WUP, the clipped image achieves high precision and time effectiveness, as depicted in Table 1. It produces nearly close to the anticipated output. When measured against the raw data, our WUP

algorithm has outperformed it in terms of accuracy and time economy. Our WUP algorithm attained the following goals:1) Greater precision; 2) time management; 3) Providing preliminary data to improve and streamline WUP; 4) a Multi-view dataset for the recognition of elephants significantly; and 5) fine-tuned back-propagation.

Table 1: Comparative analysis

Experimentation (Testing)	Single view		Multi-view	
	Accuracy	Time (s)	Accuracy	Time (s)
1	95.5	176	96.5	165
2	96.5	138	96.7	120
3	96.5	60	98.5	60
4	96.1	60	97.9	50
5	94.5	160	95.9	160
6	98.5	63	98.5	110
7	98.5	70	98.5	145
8	98.6	55	98.5	120
9	98.5	65	97.8	110
10	98	70	98	140

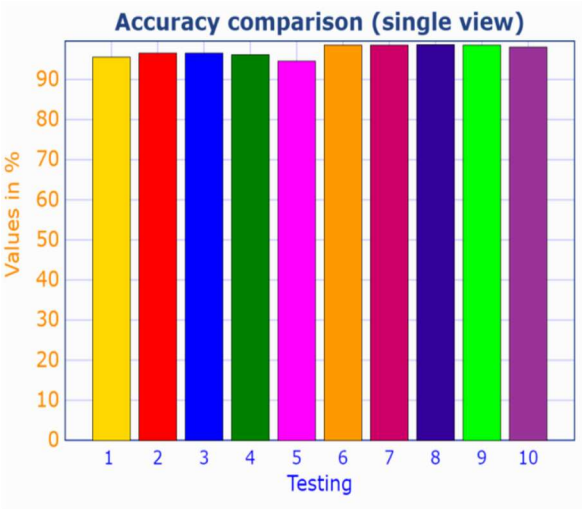

Figure 4a: Accuracy comparison for single view

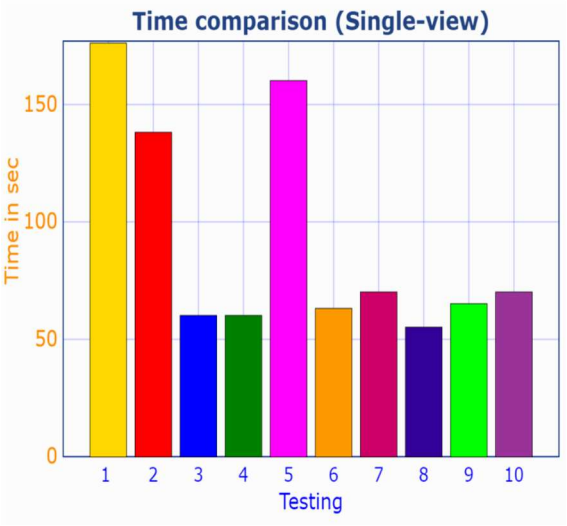

Figure 4d: Time comparison for single view

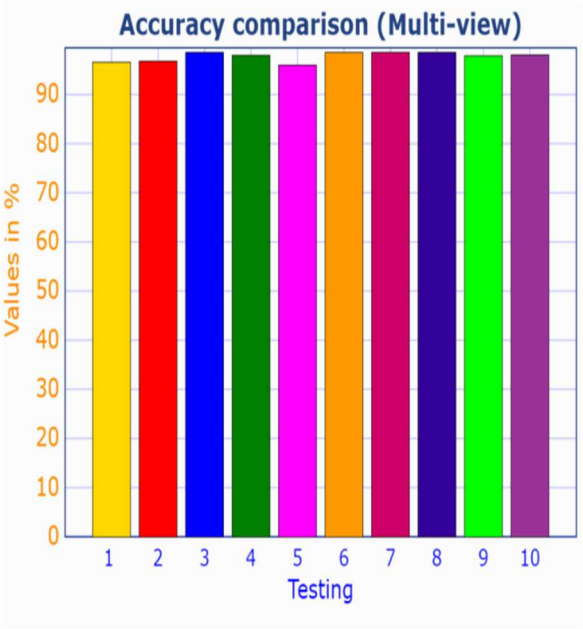

Figure 4c: Accuracy comparison for multi view

Journal of Theoretical and Applied Information Technology
15th February 2024. Vol.102. No 3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

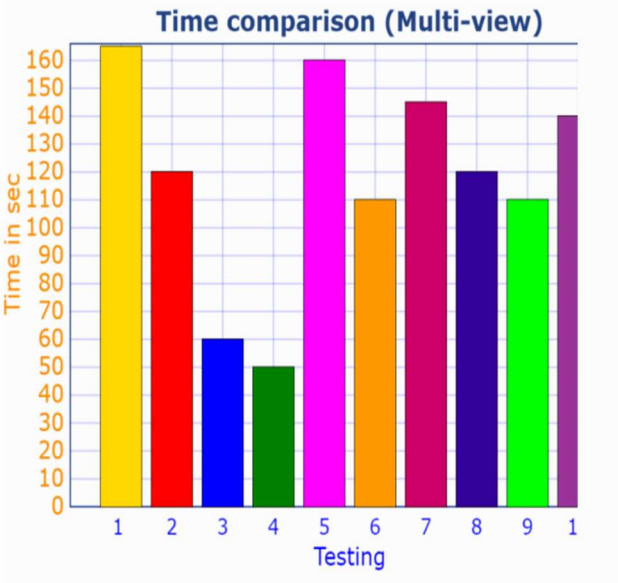

Figure 4d: Time comparison for multi view

The primary performance metric to assess the WUP algorithm's performance is accuracy. The amount of WUP iterations directly affects how accurate a measurement is. Accurate results are largely achieved through iteration. However, iteration is time-consuming even when the best methods are used, such as back-propagation and classic gradient descent. Therefore, the WUP method introduces the preparatory step, i.e., the cropping module, to reduce the number of repeats. By avoiding extraneous pixels, the cropped image requires less iteration to accurately train and assess the object that is present in it. As a result, our WUP algorithm delivers improved accuracy with less iteration. A lower number of iterations also save time. Reduced false-negative and false-positive predictions follow higher accuracy. We now consider the input photos used in our test experiments to compute various performance metrics. Twenty-five non-elephant photos and 175 elephant photographs were used in our research are considered. Our WUP algorithm produced twenty-two negative predictions and 170 positive predictions. Positive forecasts, in this case, refer to elephants, whereas negative predictions refer to others. As a result, other DL algorithms produce the predictions indicated in Fig. 5 to 9. accuracy can be calculated as follows:

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (6)$$

$$False\ positive = \frac{FP}{TN+FP} \quad (7)$$

$$False\ Negative = \frac{FN}{TP+FN} \quad (8)$$

Elephant movement generates seismic signals, which are analyzed using parameters like filter bands, altering distance from sensor and processing window size. While the annotated ground facts of elephants presence are considered genuine events, recorded background noises are categorized as false events. True positive (TP) refers to when the detection method properly identifies an existing event, while a false negative (FN) refers to all missing details from the actual event. Similarly, if algorithms successfully miss erroneous occurrences, It is a true negative (TN). However, if they mistakenly identify a false event as genuine, It is known as a false positive (FP). The results of detection methods are displayed in Fig 5 to Fig 8 as the FN, TP, FP, and TN on a time-series data. It shows the fake alert in line with the documented background sound and missed and detected occurrences for WUP on the data collected of elephant motions. As illustrated in Fig. 5 to Fig 8, Eq. 6 to Eq. 7 are used to determine the performance measures accuracy, F1-score, TPR, PPV, and FPR for the elephant movement's seismic signature that was recorded.

Table 2: Comparative analysis

Methods	TP	TN	FP	FN	Accuracy
GNet	154	15	12	5.6	87.5%
YoloV3	157	16	10.5	4.9	90.5%
ResNEt-50	158	20	10	3.2	90.7%
ResNet-150	162	21	8	3.5	93%
VGG-16	168	21	5	2.6	95.5%
VGG16+DN	170	23	2.5	1.5	97.5%
AlexNet	171	24	3	2.6	98%
WUP	175	25	2	1.1	99.5%

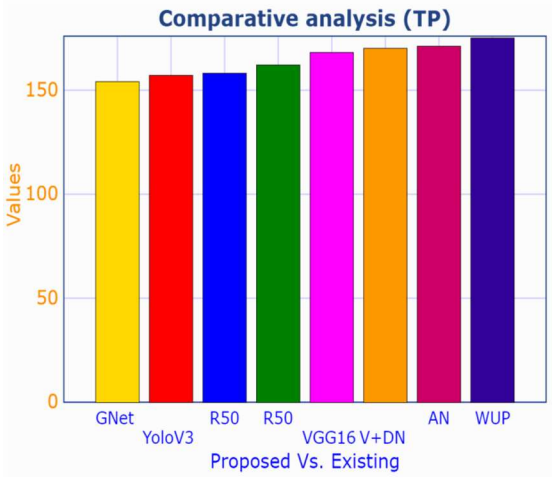

Figure 5: TP comparison

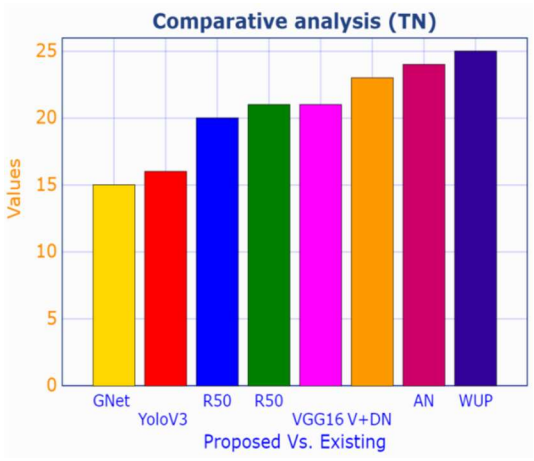

Figure 6: TN comparison

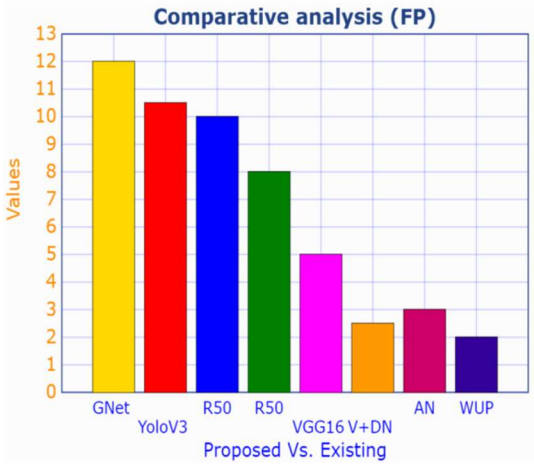

Figure 7: FP comparison

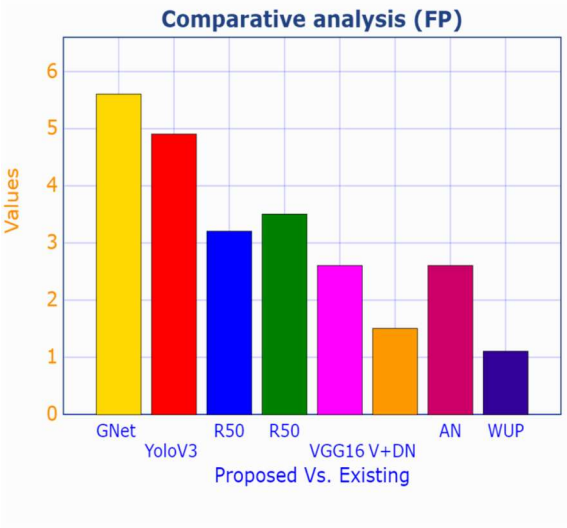

Figure 8: FP comparison

When computing FP, TN, FN, and TP, the signal is labelled as 1, which indicates the presence of elephant footsteps, and 0, which indicates background noise. The average time-series data is determined for every window as it passes over specified regions. Event is deemed to have been discovered and labeled as 1 for detected occurrences. For missed occurrences, it is 0 when energy content of time-changing window exceeds pre-defined amount. Similarly, time-series background sound is indicated as 0 for the missing event and 1 for the incorrectly detected event. While computing using, FP and TN were roughly determined using ground truth 0 and 1 for background noise and elephant motion. However, TP and FP decrease as the detection threshold increases. The appropriate threshold was selected to create a just exchange between FP and TP.

4.1. Discussion

The evaluation metric of various algorithms for varying factors is mostly covered in this section. The corner frequencies of the filter bands utilized to screen the seismic signal were 10-30 Hz, considering the dominating information on frequency provided. The algorithms are used, and for each filter band, the performance characteristics are determined for the specified window size. Moreover, the approximate amount of filter bands for comparing the performance of specified window size. Eq. 6 to Eq. 8 is used to derive the performance metrics FPR for various prevailing

approaches. A pair of elephants and one elephant alone have the maximum TPR, corresponding to a 1 s window size. Increased window lengths have resulted in a fall in FPR and an increase in TPR. It does not significantly alter as the window duration increases beyond 1s. Fig 5 to Fig 8 illustrates how FPR and TPR for the specified threshold change as a function of window size (d). However, as the detection threshold was raised, the FPR and TPR fell, affecting the F1 score. In order to ensure a reasonable trade-off between the TPR and FPR, a window size of 1s is chosen as the ideal value. Fig 9 in the WUP displays the TPR's performance with a time-varying window. Because the FPR for both a single and a pair of elephants was lowest for similar window length, maximum F1-score is obtained with a window size of 0.26s. As the window length is increased for an empirical threshold of 0.13, both the FPR and TPR increase, and the score of F1 subsequently drops. Therefore, 0.26s window length has been considered best for WUP. Although the TPR for a window size of 0.25s is the smallest, the best F1-score is created for the same window size by the lowest FPR. The FPR and TPR increase as the window length increases. Therefore, the 0.25 s ideal window length has been taken into account.

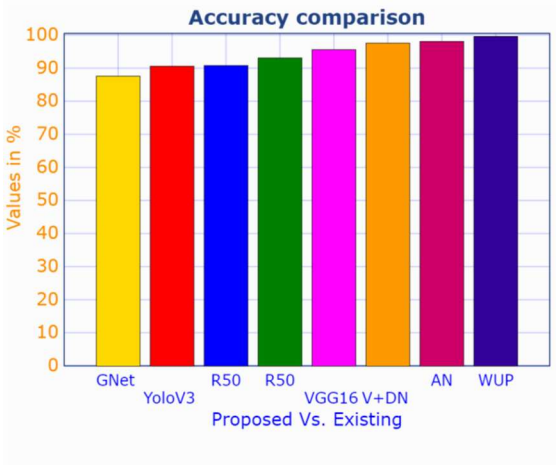

Figure 9: Accuracy comparison

4.2. Analysis

Compared to other DL algorithms, our WUP approach achieves higher accuracy rates. The performance of the WUP algorithm was assessed using various effective DL algorithms (pre-trained). FP, FN and accuracy have produced significant variations and are close to the anticipated output. The main goal of this work is to identify elephants using cameras at a false boundary (between populated areas and a forest border). Camera positioning is crucial to capture something coming into the frame. During our investigation, we tested three alternative camera angle perspectives.

Scenario 1: The camera is fetched at 0° from ground level. Considering the camera's field of view as 180°, this viewpoint can only cover 90° of that. Therefore, 0° fetching works best for catching small species rather than our intended objective, elephants.

Scenario 2: To capture a whole 180 of an object approaching the scene, angle the camera at 45 degrees. Despite the bigger target species, our test trials captured only a portion of the image. Therefore, 45° fetching is also inappropriate.

Scenario 3: In the 90° position with the camera, the full image of a grown elephant approaching the sensor unit may be seen at this precise angle. The camera is positioned at a 90-degree angle in a light pole in our pseudo boundary.

The spacing between the subject and the camera determines the image's clarity. The degree of zooming and light output is key factors in producing an image with good clarity. Test trials were carried out within the sensor unit to determine the threshold level with items positioned at various distances. The forecast is unknown when the object is placed close to the camera. Only the elephant's skin color can be seen in close-up photographs during testing. Unreliable predictions are also caused by the camera's flash failing to illuminate an object when it is too far from the camera (out-of-focus). In order to provide adequate clearance for the item captured, the threshold value has been set in 3 distinct levels, namely, medium, long, and short.

- The image appears small at a distance and should be regarded as $O(n-1)$.
- The image appears in the correct size in the middle of the range. Thus, it should be regarded as $O(n)$.
- When viewed up close, the image appears enormous but is $O(n+1)$.

Here is a close-up photo of the subject nearer to $O(n+1)$. An object is said to be out of focus in an image if located much further from $O(n-1)$. The camera is positioned at a 90-degree angle to record the object as it enters the frame from three separate angles. The screenshots are produced using two possible scenarios. (1) The camera is activated when the sensor unit detects an object. The backdrop frame is taken into account, and an image of the background is manually taken, then positioned as the opening frame. The moment a newer object enters the frame, the camera immediately takes the next frame, which is designated as frame +1. (2). Automatic photo-taking is built into the camera at the end of each period. Assume that frame is the manually set backdrop frame as δ, and even if nothing is in the frame, frame δ+1 is taken when the time is called for. Condition 1 captures everything that enters the scene even if the time frame has not yet come. Although an object does not enter the screen, Condition 2 records the scene after every designated period. The camera is placed inside the sensor unit's pseudo boundary under these two contradictory circumstances. The act of constantly capturing the elephant in movement is tiresome. There is an evaluation phase after the capture, and obtaining the image's clarity is challenging. It has already been investigated how far away each backdrop frame is. The clarity and color scheme of each background is carefully considered. The camera is actuated and takes image when one new object is present or is detected. Anticipatory models offer valuable insights for decision-making in conservation. For instance, forecasting potential conflict zones where elephants and human activities overlap enables conservationists to proactively implement measures to mitigate conflicts. Deep learning models have the capability to autonomously handle and derive pertinent features from varied data sources, including satellite imagery, GPS tracking, climate data, and others. The anticipated model can be utilized for the analysis of images, extracting valuable information from aerial or satellite imagery to comprehend the environmental conditions and terrain that might impact the movements of elephants.

5. CONCLUSION

This study characterizes the movements of elephants in the forest using a coarser-level WUP approach to alert and protects them. Data was generated repeatedly for an individual elephant and a pair of elephants, and a dataset was formulated. The sensors detect the ground vibrations caused by an elephant moving about. To establish the truth for the recording of data, camera with time synchronized and GPS are considered. Elephant locomotion's seismic signal is assessed to determine the main frequency between 0 and 20 meters and 20 and 40 meters from the sensor. Signal detection algorithms like WUP were applied for the feasibility analysis of detection on the data. Compared to various prevailing approaches, WUP provides the best general detection rate of 99.5% at radial distance of 20m from sensor with an enhanced F1–score of 30% and 15%, and 3% for 50 m. For filter bands with 20 Hz, these algorithms beat competitors. The efficient recognition of distinct elephant footstep patterns resulting from pachyderm locomotion can also be accomplished using a finer elephant detection strategy. The major research constraint is the lack of modern technical analysis like the adoption of deep learning. The evolution of deep learning gives huge research impact with advanced enhancements in prediction outcomes with reduced computational complexity. The overhead due to the available samples, i.e. under-sampling which is encountered in the anticipated model is reduced in the future.

Conflicts of Interest The authors declare no conflict of interest. Author Contributions The paper background work, conceptualization, methodology, dataset collection, implementation, result analysis and comparison, preparing and editing draft, visualization have been done by first author. The supervision, review of work and project

REFERENCES:

[1] S. Blake, S. Strindberg, P. Boudjan, C. Makombo, I. Bila-Isia, O. Ilambu, et al., Forest elephant crisis in the Congo Basin, *PLoS Biology*, 5(4), 2007, e111.

[2] J. R. Poulsen, S. E. Koerner, S. Moore, V. P. Medjibe, S. Blake, C. J. Clark, et al., Poaching empties critical Central African wilderness of forest elephants, *Current Biology*, 27(4), 2017, R134–R135.

[3] S. Albelwi and A. Mahmood, Analysis of instance selection algorithms on large datasets with deep convolutional neural networks. *IEEE Long Island Systems,*

Applications and Technology Conference (LISAT), NY, USA, 2016, pp. 1-5.

[4] S. Albelwi and A. Mahmood, Analysis of instance selection algorithms on large datasets with deep convolutional neural networks. *IEEE Long Island Systems, Applications and Technology Conference (LISAT)*, NY, USA, 2016, pp. 1-5.

[5] J. S. Anni, and A. K. Sangaiah, Wireless integrated sensor network: boundary intellect system for elephant detection via cognitive theory and fuzzy cognitive maps, *Future Generation Computer Systems*, 83 (2018), 522–534.

[6] P. Christiansen, L. N. Nielsen, K. A. Steen, R. N. Jrgensen, and H. Karstoft, Deep anomaly: combining background subtraction and deep learning for detecting obstacles and anomalies in an agricultural field, *Sensors*, 16(1904), 2016, 1–21.

[7] X. Gengjian, S. Jun, and S. Li, Background subtraction based on phase feature and distance transform, *Pattern Recognition Letters*, 33(12), 2012, 1601–1613.

[8] S. Hang, S Maji, E. Kalogeraskis and L. M. Erik, Multi-view convolutional neural networks for 3D shape recognition, *Proceedings of the IEEE International Conference on Computer vision*, Santiago, Chile, 2015, pp. 945–953.

[9] F. Hu, G. S. Xia, J. Hu and L. Zhang, Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery, *Remote Sensing*, 7(11), 2015, 14680–14707.

[10] J. Li, L. Sun, Q. Yan, Z. Li, W. Srisa- An, and Y. Yeng, Significant permission identification for machine learning based android malware detection, *IEEE Transactions on Industrial Informatics*, 14(7), 2018, 3216–3225.

[11] L. Ping, L. Jin, H. Zhenga, L. Tong, Z. G. Chong, M. Y. Siu and K. Chen, Multi-key privacy-preserving deep learning in cloud computing, *Future Generation Computer Systems*, 74 (2017), 76–85.

[12] H. Yang, K. Sirlantzis, G. Howells, N. Ragot, and P. Rodrguez, An online background subtraction algorithm deployed on a NAO humanoid robot based monitoring system, *Robotics and Autonomous Systems*, 85 (2016), 37–47.

[13] M. Zeppelzauer and A. S. Stoeger, Establishing the fundamentals for an elephant early warning and monitoring system, *BMC Research Notes*, 8(409), 2015, 1–15.

[14] X. Zheng, L. Xiong, W. Le, Y. Qiuwei, D. Jiayi and A. K. Sangaiah, Using convolution control block for Chinese sentiment analysis, *Journal of Parallel and Distributed Computing*, 116 (2018), 18–26.

[15] P. Sherubha,S. P. Sasirekha, V. Manikandan, K. Gowsic and N. Mohanasundaram, Graph Based Event Measurement for Analyzing Distributed Anomalies in Sensor Networks, *Sādhanā*, 45 (2020), 1-5.

[16] P. Sherubha, An Efficient Network Threat Detection and Classification Method using ANP-MVPS Algorithm in Wireless Sensor Networks, *International Journal of Innovative Technology and Exploring Engineering*, 8(11), 2019, 1-5.

[17] P. Sherubha, An Efficient Intrusion Detection and Authentication Mechanism for Detecting Clone Attack in Wireless Sensor Networks, *Journal of Advanced Research in Dynamical and Control Systems*, 11(5), 2019, 55-68.

[18] H. Zhou, Y. Chen, and R. Feng, A novel background subtraction method based on color invariants, *Computer Vision and Image Understanding*, 117(11), 2013, 1589–1597.

[19] A. Akula, R. Ghosh, S. Kumar, and H. K. Sardana, Moving target detection in thermal infrared imagery using spatiotemporal information, *JOSA A*, 30(8), 2013, 1492–1501.

[20] J. S. Anni and A. K. Sangaiah, Wireless Integrated Sensor Network: Boundary Intellect system for elephant detection via cognitive theory and Fuzzy Cognitive Maps, *Future Generation Computer Systems*, 83 (2018), 522–534.

[21] V. Duk, L. Rosenberg, and B. W. H. Ng, Target detection in sea-clutter using stationary wavelet transforms, *IEEE Trans. Aerosp. Electron. Syst*. 53 (2017), 1136–1146.

[22] R. Ghosh, A. Akula, S. Kumar, and H. K. Sardana, Time-frequency analysis based robust vehicle detection using seismic sensor, *Journal of Sound and Vibration*, 346 (2015), 424–434.

[23] R. Ghosh, A. Vajpeyi, A. Akula, V. Shaw, S. Kumar, and H. K. Sardana, Performance evaluation of a real-time seismic detection

system based on CFAR detectors, *IEEE Sensors Journal*, 20 (2020), 3678–3686.

[24] X. Jin, S. Sarkar, A. Ray, S. Gupta, and T. Damarla, Target detection and classification using seismic and PIR sensors, *IEEE Sensors Journal*, 12 (2012), 1709–1718.

[25] C. E. O'Connell-Rodwell, B.T. Arnason, and L.A. Hart, Seismic properties of Asian elephant (Elephas maximus) vocalizations and locomotion, *Journal of Acoustic Society of America*, 108 (2000), 3066–3072.

[26] C. E. O'Connell-Rodwell, L. A. Hart, and B.T. Arnason, Exploring the potential use of seismic waves as a communication channel by elephants and other large mammals, *American Zoologist*, 41 (2001), 1157–1170.

[27] J. D. Wood, C.E. O'Connell-Rodwell, S.L. Klemperer, Using seismic sensors to detect elephants and other large mammals: a potential census technique, *Journal of Applied Ecology,* 42 (2005), 587–594.

[28] G. Wang, Machine learning for inferring animal behavior from location and movement data, *Ecological Informatics*, 49 (2019), 69–76.

[29] S. J. Sugumar, R. Jayaparvathy, An improved real time image detection system for elephant intrusion along the forest border areas, *Current Science*, 104 (2013), 1515–1526.

[30] C. H. Lin, K.C. Hsu, K.R. Johnson, et al., Evaluation of machine learning methods to stroke outcome prediction using a nationwide disease registry, *Computer Methods and Programs in Biomedicine*, 190(2020), 105381.