Identification of Wild Animals in Forest Surveillance Cameras

Prathyusha Dokku(✉), Swapna Mudrakola, Kalyan Kumar Dadi, and Nikhitha Akula

Department of Computer Science and Engineering, Matrusri Engineering College, Hyderabad 500059, India
cse19733083@matrusri.edu.in

Abstract. In the ever-expanding realm of wildlife conservation and ecological research, the use of automated image classification software has emerged as a valuable tool for extracting crucial insights from camera trap images. However, a persistent challenge lies in the software's ability to maintain consistent performance and spatial independence for a given image, thus necessitating a solution to enhance its location invariance. The paper introduces an optimized location-invariant camera trap object detector, trained with publicly available image datasets, demonstrating a significant performance improvement with an epoch accuracy of up to 99%. This innovative approach not only addresses the current limitations but also opens avenues for more robust and globally applicable wildlife monitoring solutions, fostering advancements in ecological understanding and conservation efforts.

Keywords: Neural Networks · VGG · Image processing techniques

1 Introduction

The unbalanced ecosystem and modern lifestyle have resulted in two major problems: fire accidents and wildlife crimes. Deforestation is the primary cause of human threat to people living around the boundaries of the forest and small villages around the forest. This has led to wild animals invading villages in search of food and water, putting the lives of people living in those areas at risk. For instance, in Kombukuthi village, nearly 430 families are living in danger zones [1]. The reasons for the locomotion of wild animals to villages and panchayats are illegal mining, deficiency of food, water, shelter, and deforestation, and Keonjhar Sardar Forest is among them [2]. Forest fires are one of the major problems for the world's ecosystem and are inevitable due to changes in temperature [3]. Throughout the world, there is a threat of epidemic plant life loss [4]. As per the survey report of State of Forests Report, 2021 (SoFR, 2021) India, which was released on Jan 2022, there is an increase in the overall number of forest fire attacks recorded which is 345,989 from Nov 2020 to June 2021. Forest fire attacks were very low in comparison to the current situation [5]. There are

© ICST Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 2024
Published by Springer Nature Switzerland AG 2024. All Rights Reserved
O. Castillo et al. (Eds.): PerSOM 2023, LNICST 517, pp. 229–239, 2024.
https://doi.org/10.1007/978-3-031-66044-3_16

many reasons for animals invading the village like illegal poaching, smuggling of epidemic trees, and low survival once maintenance by the Forest department is common reasons across the world [6].

Machine learning techniques have found a wide array of applications across diverse domains, showcasing their adaptability and effectiveness. In the realm of sentiment analysis in e-commerce customer reviews [28], methods for sentiment detection have been significantly enhanced, providing businesses with invaluable insights into customer satisfaction and feedback. Moreover, the utilization of AI in the detection of impersonators in examination halls [29] addresses the critical need for maintaining the integrity of educational assessments. Additionally, multiple face detection, employing algorithms like Haar-AdaBoosting, LBP-AdaBoosting, and neural networks [30], demonstrates the versatility of machine learning in enhancing security and surveillance systems. These instances underscore the pivotal role that machine learning plays in solving real-world challenges through surveillance and monitoring.

Sensors have versatile applications, including temperature and motion detection [7]. An innovative use is estimating gas leakage rates in extensive pipeline systems [27] via a fuzzy interface algorithm. These sensors are placed in various devices and environments to collect data, which is then transmitted wirelessly, enhancing efficiency [8,9].

Artificial Intelligence (AI) further empowers IoT by simulating human-like intelligence [10]. AI finds applications in facial recognition, automation, and chatbots, leveraging technologies like Deep Learning, Cloud Computing, and Quantum Computing. It automates processes across various sectors [11].

IoT and AI combined give rise to voice assistants and smart devices, enhancing computer-human interactions and analytics like Deep Learning, Data Analytics, and Machine Learning [12,13]. Sensors integrated with AI have notably improved analytics and alert systems [14]. In email alert systems [15], they enable real-time communication and responsive alert mechanisms, enhancing monitoring and communication.

Automated image classification [23] plays a pivotal role in the realm of wildlife monitoring and management. It provides a sophisticated means of efficiently processing vast quantities of camera trap images, offering insights into wildlife populations and behaviors. However, the challenge of location variability can hinder the effectiveness of such systems. The development of location-invariant [22] object detectors emerges as a crucial endeavor to overcome this challenge. These detectors aim to ensure consistent and accurate classification of animals in diverse environmental settings, regardless of factors such as lighting conditions and camera angles. The primary aim of this study is to explore the possibility of using publicly available datasets for training location-invariant models. By evaluating the adaptability and transferability of these models to the unique context of camera trap images, our research endeavors to expand our comprehension of their capacity to augment the consistency and effectiveness of automated image classification.

2 Literature Survey

Object detection is a crucial task in computer vision that involves identifying and localizing objects within images or videos. Various algorithms and techniques have been proposed to address this challenge. One such study by Fares Jalled et al. [15] utilized image processing techniques, including feature detection and the Haar detection cascade, to detect specific objects such as human bodies and cars. Their algorithm demonstrated effective object detection even with changes in scale or minor plan rotations. Similarly, Zeyad Al-Zaydi et al. [16] focused on detecting and counting people in crowded scenes using the GMM algorithm and a trained GRP regression model. Their approach showed promise in applications such as abnormality detection and crowd control.

The deep-learning architecture, utilizing CNN, Bi-LSTM, and attention mechanisms, has demonstrated significant promise in accurately categorizing and detecting aberrant human behavior in video streams enhancing the wide applicaiton of Human behaviour monitoring systems [26]. In the context of wildlife monitoring, N. Banupriya et al. [17] employed Convolutional Neural Networks (CNN) for accurate and efficient animal detection. A. W. D. Udaya Shalika et al. [18] developed an animal classification system using camera/PIR detection and machine learning with SVM, achieving an overall accuracy of 80%. Falzon et al. [19] explored location-invariant animal detection in camera trap images using publicly available sources and achieved promising results in mean Average Precision (mAP). These studies collectively contribute to the fields of object detection and animal classification, offering insights and advancements in computer vision applications. The summarized information can be found in Table 1.

In the domain of camera trap image processing, prior research has made use of object detector and image classifier models, including contributions from Norouzzadeh et al. [24], and Tabak et al. [25]. These models rely on data-driven deep learning techniques for image classification, eliminating the need for manual feature engineering. The existing approach struggles to adapt to images from various places, needs a lot of resources to make location-specific models, and faces difficulties when specialized models aren't accessible. This study tackles these problems by concentrating on improving location-flexible camera trap object detectors. The proposed methodology explores training models using publicly available image datasets.

3 Methodology

A camera trap, equipped with a PIR sensor, is designed for wildlife image capture [21]. The PIR sensor comprises two vital components: the emitter projecting an infrared beam and the receiver detecting this beam. It operates by allowing the emitted infrared light to pass through the environment undisturbed. However, when an obstruction, often an animal, disrupts the light path, the PIR sensor registers this disturbance. For object classification of images, particularly distinguishing animals and humans, deep learning techniques are highly effective

Table 1. Summary of Object Detection Algorithms

S.No	Author	Title	Methodology	Result	Remark
1	Fares Jalled et al. [15]	Object detection using image processing with the help of UAB	Object detection using image processing with the help of UAB, which involves image capture, feature detection, collecting putative points, and object detection.	The system effectively detects specific objects like human bodies and cars. It achieves a 1% false negative and 40% false positive rate for face detection	The system performs well in detecting objects despite changes in scale or minor plan rotations. Enhanced dataset could further reduce false alarms.
2	Zeyad Al-Zaydi et al. [16]	Detecting people and counting them in the crowd.	Detecting and counting people in crowds using frames from a benchmark dataset, Gaussian Mixture Model (GMM) for background subtraction, and a trained GRP regression model.	Mean Absolute Error (MAE) is 1.945, and Mean Squared Error is 6.056. The error obtained is less, nearly 0.2.	The proposed system's results are promising and can be valuable for applications like crowd control and evacuation planning.
3	N. Banupriya et al. [17]	Animal detection using deep learning algorithm	Animal detection using a Convolutional Neural Network (CNN).	TAchieves an accuracy of 99.10%, with an epoch value of 0.8679 and a processing time of 52.208 ms.	The CNN-based system demonstrates high accuracy in animal detection.
4	A. W. D. Udaya Shalika et al. [18]	Animal Classification System - for animal researchers and wildlife photographers.	Animal classification system involving camera/PIR detection, motion and tracking, feature extraction, recognition, and machine learning (ML) using Support Vector Machine (SVM).	Overall accuracy is 80%. However, there are challenges when combining certain descriptors.	The system struggles to separate negative images, especially when combining specific descriptors.
5	Falzon et al. [19]	Automated location invariant animal detection in camera trap images using publicly available data sources	Automated location-invariant animal detection in camera trap images using publicly available data sources and an optimization strategy called 'infusion.'	Achieves mean Average Precision (mAP) results ranging from 38.5% to 88.59%.	While the approach performs well, it faces limitations in achieving high precision for out-of-sample object detection. It should be evaluated with alternative object detection frameworks and on a broader multi-class dataset.

[20]. Using convolutional neural networks (CNNs), these methods process image inputs through layers, starting with an input layer collecting all images. The hidden layer, consisting of a convolutional layer and a fully connected dense layer, extracts features from input images and assigns importance to these features. The output layer generates results for user application.

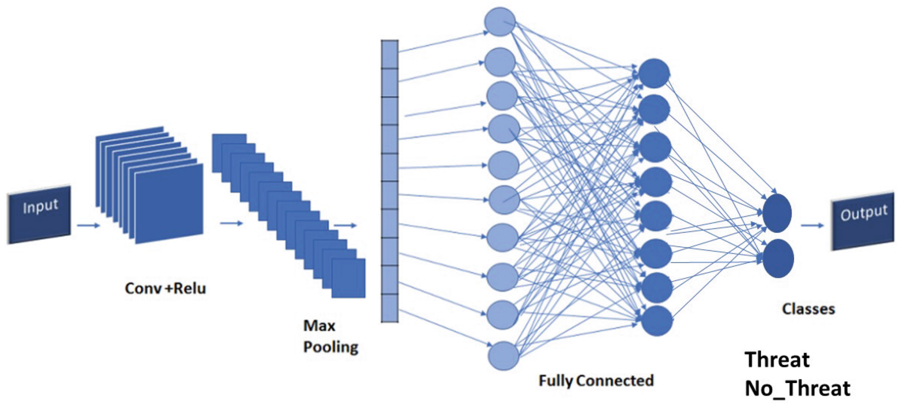


Fig. 1. System Architecture of the proposed model

The Proposed system architecture given in Fig. 1, is a simple Sequential model consisting of three layers: a Flatten layer, two Dense layers, and two Dropout layers. The Flatten layer takes the output of the last convolutional layer of the VGG16 model and flattens it into a one-dimensional array. This output is then passed on to the first Dense layer, which has 100 neurons and uses the Rectified Linear Unit (ReLU) activation function. The purpose of this layer is to learn higher-level features by combining the information from the flattened output of the previous layer. Next, a Dropout layer is applied, which randomly drops out some of the neurons during each training epoch to prevent overfitting. This is followed by a second Dense layer with 50 neurons and again using the ReLU activation function. The purpose of this layer is to further refine the learned features. To handle the location invariance, we have added a GAP (Global Average Pooling) layer to enhance the model's performance in recognizing objects across different positions and orientations within the images. Finally, another Dropout layer is applied, followed by the last Dense layer with 2 neurons and the Softmax activation function.

The Softmax function outputs a probability distribution over the two classes - threat and non-threat. This architecture is designed to classify the given image as either a threat or non-threat based on the learned features. The initial dataset used for training the proposed system consisted of 15,000 images sourced from publicly available repositories such as iNaturalist. After a rigorous data cleaning process, the dataset was refined to include 7,546 images for the training phase and an additional 1,574 images for validation. The trained model was subsequently tested using a dataset comprising 3,732 images to assess its performance and obtain the desired output.

3.1 Algorithm for the Proposed System:

1. Import necessary libraries
   - `import keras`
   - `import numpy`
   - `import pandas`
   - `import matplotlib`
2. Load and preprocess the dataset
   - `dataset = load_dataset()`
   - `train_data, validation_data = split_data(dataset)`
   - `preprocess(train_data)`
   - `preprocess(validation_data)`
3. Define the model
   - `model = Sequential()`
4. Add layers to the model
   - `model.add(Flatten())`
   - `model.add(Dense(100, activation='LeakyReLU')`
   - `model.add(Dropout(0.5))`
   - `model.add(Dense(50, activation='LeakyReLU'))`
   - `model.add(Dropout(0.3))`
   - `model.add(GlobalAveragePooling2D())`
   - `model.add(Dense(output_classes, activation='softmax'))`
5. Compile the model
   - model.compile(loss='categorical_crossentropy', optimizer='RMSprop', metrics=['accuracy'])
6. Train the model
   - model.fit(train_data, epochs=num_epochs, batch_size=batch_size, validation_data=validation_data)
7. Evaluate the model
   - `test_data = load_test_data()`
   - `loss, accuracy = model.evaluate(test_data)`
8. Make predictions
   - `new_data = load_new_data()`
   - `predictions = model.predict(new_data)`

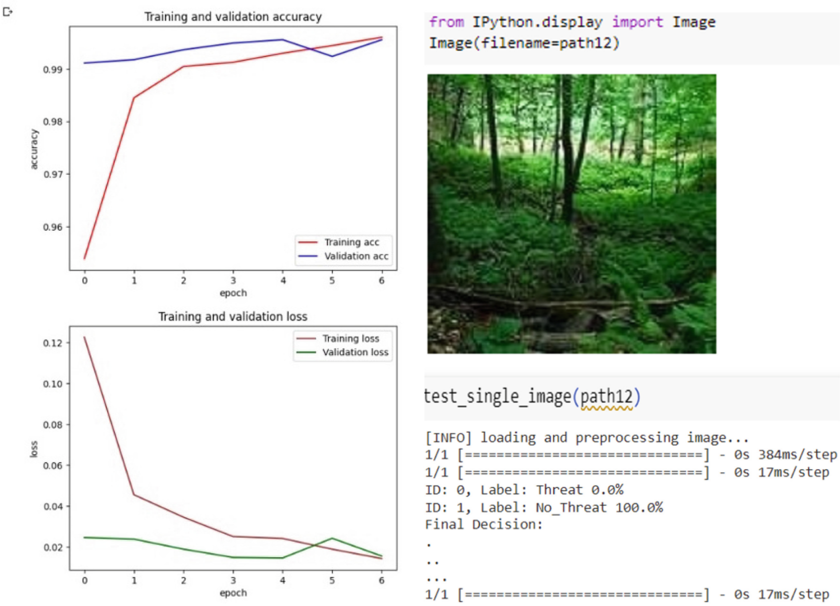
3.2 Objectives and Advantages of the Proposed System

The proposed system aims to address critical challenges in wildlife monitoring, ecosystem protection, and human-wildlife interaction.

1. Objectives
   - Automated Image Classification: Develop an automated image classification system using Convolutional Neural Networks (CNN) to distinguish between animals and humans in camera trap images.
   - Creating location-invariant object detectors to improve image classification accuracy, especially in the context of camera traps.
2. Advantages
   - Safeguard agricultural fields from wildlife intrusion, reducing crop damage.
   - Contribute to the conservation of forest ecosystems and wildlife by monitoring and addressing potential threats.
   - Enhance the management of zoological parks and wildlife sanctuaries by automating animal monitoring.
   - Improve safety and enhance the experience of tourists and researchers in areas with high human-wildlife interactions, such as safaris.
   - Minimize risks for trekkers and hikers in regions with known wildlife presence.
   - Serve as a deterrent to illegal activities like animal poaching and provide evidence for law enforcement.
   - Facilitate the observation and study of animal behavior in their natural habitats.
   - Enhance the protection of tribal and local communities living near wildlife habitats.
   - Automatically notify relevant authorities, including forest departments, village leaders, and local police, in case of potential threats or incidents.

4 Results and Discussions

A holistic approach to mitigate animal threats in forest areas benefits both human and animal safety, preventing illegal poaching. The integration of Global Average Pooling (GAP) in the model has significantly enhanced its ability to recognize objects in diverse locations and orientations. The model exhibits a robust precision of 0.99, accurately identifying 99% of cases, with a recall of 0.98 (Table 2). Figure 2 presents training and validation accuracy as well as loss plots using the matplotlib library, providing insights into the machine learning model's performance. Further refinements are possible for increased reliability, particularly in identifying threats with varied object positions within images.



(a) Caption for Image 1 (b) Caption for Image 2

Fig. 2. Images of accuracy and sample test case

Table 2. Model Performance Metrics

Class	Precision	Recall	F1-Score	Support
Threat	1.00	1.00	1.00	2961
No Threat	0.99	0.98	0.98	771
Micro-Avg	0.99	0.99	0.99	3732
Macro-Avg	0.99	0.99	0.99	3732
Weighted-Avg	0.99	0.99	0.99	3732
Samples-Avg	0.99	0.99	0.99	3732

The model effectively distinguishes between "Threat" and "No_Threat" instances depicted in Fig. 3, with 2950 "threat" instances correctly classified. There are minimal misclassifications: 11 "Threat" instances are incorrectly labeled as "No_Threat," and 13 "No_Threat" instances are misclassified as "threat." These minor errors indicate potential areas for improvement. Furthermore, 758 "No_Threat" instances are accurately identified, showcasing the model's proficiency in recognizing instances without threats.

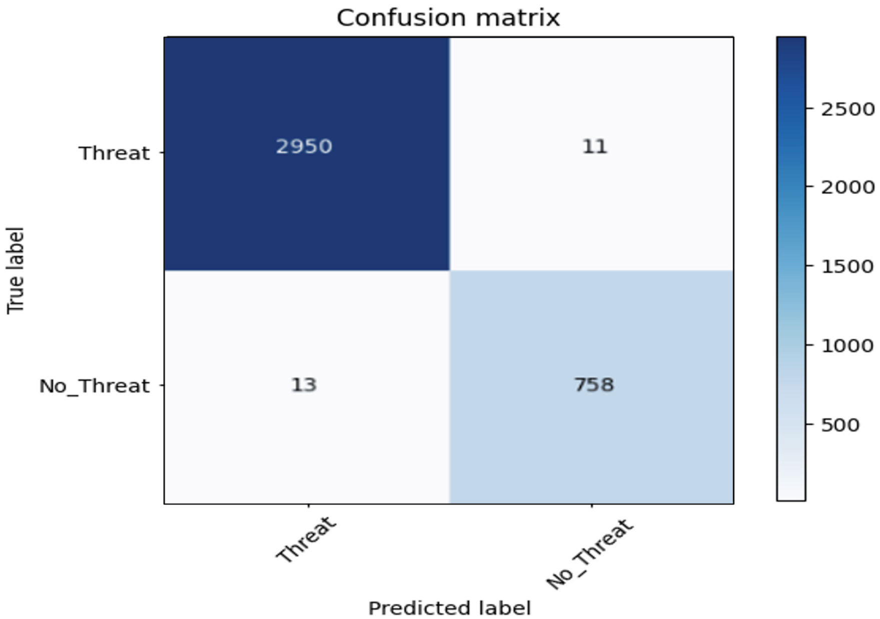


Fig. 3. Performance Matrix of Test Data

5 Conclusion

We have used CNN model which gave higher precision and accuracy compared to other models and there is still room for development of more profound location invariant models. The main advantage of implementing this proposed system generates less intervention of humans in forests and also we can get 24/7 surveillance without issues of exhaustion or reduction in quality of work. The development of an alert system to prevent animal threats is possible and there is a further scope of advancement in the classification which is now to animals or humans and can be used for fire detection or any other forest problems. There is a scope of using sound sensors for detection of the emotions of animals near the region border whether calm or scared causing trouble or harm.

6 Future Scope

The envisioned system holds potential for extension into reptile detection and the identification of diverse fauna. A promising avenue involves developing an expert system for dynamic, real-time analysis within specific forest areas. This expansion not only enhances the system's versatility but also offers a more comprehensive approach to ecological studies. The envisioned expert system can adapt to evolving scenarios, making it invaluable for ongoing monitoring and analysis, thus contributing to a nuanced understanding of the dynamic ecosystems we aim to explore and preserve.

References

1. New Indian Express. "430 families in Kombukuthi tribal village live in fear of wild animal attacks". Retrieved from https://www.newindianexpress.com/states/kerala/2022/feb/06/430-families-in-kombukuthi-tribal-village-live-in-fear-of-wild-animal-attacks-2416023.html. Accessed 26 Apr 2023
2. Times of India. "Animals flee forests in search of food, water". https://timesofindia.indiatimes.com/city/bhubaneswar/animals-flee-forests-in-search-of-food-water/articleshow/13060423.cms. Accessed 26 Apr 2023
3. NASA Earth Observatory. "MOD14A1 M FIRE". https://earthobservatory.nasa.gov/global-maps/MOD14A1_M_FIRE/. Accessed 26 Apr 2023
4. Down To Earth. "Forest Survey Report 2021: Forest fire counts up 2.7 times". https://www.downtoearth.org.in/news/forests/forest-survey-report-2021-forest-fire-counts-up-2-7-times-81123. Accessed 26 Apr 2023
5. World Wildlife Fund. "Illegal Wildlife Trade". https://www.worldwildlife.org/threats/illegal-wildlife-trade. Accessed 26 Apr 2023
6. World Economic Forum. "New high-resolution map shows fires caused one-third of global forest loss between 2001 and 2019". Retrieved from https://www.weforum.org/agenda/2022/03/new-high-resolution-map-shows-fires-caused-one-third-of-global-forest-loss-between-2001-and-2019/. Accessed 26 Apr 2023
7. Ahmad, M.B., Abdullahi, A.A., Muhammad, A.S., Saleh, Y.B., Usman, U.B.: The various types of sensors used in the security alarm system. Int. J. New Comput. Archit. Appl. (IJNCAA) **9**(2), 50–59 (2019)
8. Zeinab, K.A.M., Elmustafa, S.A.A.: Internet of things applications, challenges and related future technologies. World Sci. News **67**(2), 126–148 (2017)
9. Ullo, S.L., Sinha, G.R.: Advances in smart environment monitoring systems using IoT and sensors. Sensors **20**(11), 3113 (2020)
10. Wang, P.: What Do You Mean by "AI"? In: AGI, vol. 171, pp. 362–373 (2008)
11. Smith, R.G., Eckroth, J.: Building AI applications: yesterday, today, and tomorrow. AI Mag. **38**(1), 6–22 (2017)
12. Ghosh, A., Chakraborty, D., Law, A.: Artificial intelligence in the Internet of things. CAAI Trans. Intell. Technol. **3**(4), 208–218 (2018)
13. Šumak, B., Brdnik, S., Pušnik, M.: Sensors and artificial intelligence methods and algorithms for human-computer intelligent interaction: a systematic mapping study. Sensors **22**(1), 20 (2021)
14. Leung, C.K., Braun, P., Cuzzocrea, A.: AI-based sensor information fusion for supporting deep supervised learning. Sensors **19**(6), 1345 (2019)
15. Jalled, F., Voronkov, I.: Object detection using image processing. arXiv preprint arXiv:1611.07791 (2016)
16. Al-Zaydi, Z., Vuksanovic, B., Habeeb, I.: Image processing based ambient context-aware people detection and counting. Int. J. Mach. Learn. Comput. **8**(3), 268–273 (2018)
17. Banupriya, N., Saranya, S., Swaminathan, R., Harikumar, S., Palanisamy, S.: Animal detection using a deep learning algorithm. J. Crit. Rev. **7**(1), 434–439 (2020)
18. Shalika, A.U., Seneviratne, L.: Animal classification system based on image processing & support vector machine. J. Comput. Commun. **4**(1), 12–21 (2016)
19. Shepley, A., Falzon, G., Meek, P., Kwan, P.: Automated location invariant animal detection in camera trap images using publicly available data sources. Ecol. Evol. **11**(9), 4494–4506 (2021)

20. Ranganayagi, D., Saranya, P., Sharmila, M.J., Sujitha, S., Nisha, T.A., Shanmugam, K.: Pre-eclampsia risk monitoring and alert system using machine learning and IoT (2022)
21. Welbourne, D.J., Claridge, A.W., Paull, D.J., Lambert, A.: How do passive infrared triggered camera traps operate and why does it matter? Breaking down common misconceptions. Remote Sens. Ecol. Conserv. **2**(2), 77–83 (2016)
22. Maddern, W., Stewart, A., McManus, C., Upcroft, B., Churchill, W., Newman, P.: Illumination invariant imaging: applications in robust vision-based localisation, mapping and classification for autonomous vehicles. In: Proceedings of the Visual Place Recognition in Changing Environments Workshop, IEEE International Conference on Robotics and Automation (ICRA), Hong Kong, China,vol. 2, no. 3, p. 5 (2014)
23. Sharath Kumar, Y.H., Manohar, N., Chethan, H.K.: Animal classification system: a block based approach. Procedia Comput. Sci. **45**, 336–343 (2015)
24. Mohammad Sadegh, N., et al.: Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. Proc. Nat. Acad. Sci. **115**(25), E5716–E5725 (2018)
25. Tabak, M.A., et al.: Machine learning to classify animal species in camera trap images: applications in ecology. Methods Ecol. Evol. **10**(4), 585–590 (2019)
26. Kumar, M., Patel, A.K., Biswas, M., Shitharth, S.: Attention-based bidirectional-long short-term memory for abnormal human activity detection. Sci. Rep. **13**(1), 14442 (2023)
27. Manoharan, H., Shitharth, S., Sangeetha, K., Kumar, B.P., Hedabou, M.: Detection of superfluous in channels using data fusion with wireless sensors and fuzzy interface algorithm. Meas. Sens. **23**, 100405 (2022)
28. Singh, U., Saraswat, A., Azad, H.K., Abhishek, K., Shitharth, S.: Towards improving e-commerce customer review analysis for sentiment detection. Sci. Rep. **12**(1), 21983 (2022)
29. Vishal, A., Nitish Reddy, T., Reddy, P.P., Shitharth, A.: Detecting impersonators in examination halls using AI. 344–350 (2021)
30. Devi, B.T, Shitharth, S.: Multiple Face Detection Using Haar-AdaBoosting, LBP-AdaBoosting and Neural Networks. In IOP Conference Series: Materials Science and Engineering, vol. 1042, no. 1, p. 012017. IOP Publishing (2021)