Journal of Theoretical and Applied Information Technology
15th December 2024. Vol.102. No. 23

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

DUAL Q-LEARNING WITH GRAPH NEURAL NETWORKS: A NOVEL APPROACH TO ANIMAL DETECTION IN CHALLENGING ECOSYSTEMS

JOHNWESILY CHAPPIDI [1] DIVYA MEENA SUNDARAM [2*]

[1] Research Scholar, School of Computer Science and Engineering, VIT-AP University, Andhra Pradesh, India

[2] Assistant Professor, School of Computer Science and Engineering, VIT-AP University, Andhra Pradesh, India

E-mail: [1]johnwesily.21phd7098@vitap.ac.in, [2]divyameena.s@vitap.ac.in

ABSTRACT

Detecting wild animals is crucial to prevent road accidents caused by their crossings and mitigate intrusions into residential areas. Existing methods often struggle with complex spatial contexts and environmental variability. This study introduces an integrated approach using Graph Neural Networks (GNNs), advanced Q Learning, Multi-Attribute Utility Theory (MAHP) with deep learning, and Generative Adversarial Networks (GANs) for data augmentation. The model enhances spatial awareness with GraphSAGE and Graph Attention Networks (GAT), employs Deep Q-Networks (DQN) for adaptive decision-making, integrates MAHP with custom CNNs for nuanced attribute evaluation, and utilizes Conditional GANs for synthetic data generation. Comparative evaluations show substantial enhancements in accuracy, precision, recall, speed, AUC, and specificity, establishing new benchmarks for wildlife detection in challenging conditions. This research advances automated wildlife monitoring, which is crucial for biodiversity conservation and addressing ecological challenges through integrated computational techniques.

Keywords: *Graph Neural Networks, Advanced Q Learning, Multi-Attribute Utility Theory, Generative Adversarial Networks*

1. INTRODUCTION

Wildlife detection in Complex Environments becomes crucial for biodiversity conservation, preventing Animal-human conflict (AHC), and avoiding fatalities caused by animals crossing roads unexpectedly. As human activities expand into natural areas, encounters between wildlife and human infrastructure have increased, posing substantial ecological and societal issues. In India, approximately 50,000 animal-vehicle accidents are reported annually, leading to serious human injuries, wildlife fatalities, and economic losses [1] [2]. These frightening statistics illustrate the critical need for reliable, automated systems to detect and monitor wild animals, especially in complex environments, to improve road safety, reduce conflict, and maintain biodiversity. Despite advances in object detection technology, establishing a system capable of tackling the constraints of natural ecosystems continues to be a critical research subject.

Identifying animals in their natural environments presents numerous technical obstacles that traditional detection techniques do not adequately resolve. Occlusion is a frequent problem, as animals usually remain partially or fully hidden by dense greenery, rocks, or various objects, making them challenging to detect [3]. Furthermore, the intricacy of backgrounds, including the dynamic and varied landscapes found in forests or grasslands, frequently leads to animals merging with their environments, elevating the rates of false positives and false negatives [4]. Variations in the natural environment, including changes in illumination. Climate conditions and terrain introduce challenges to the generalization capabilities of detection models [5]. The challenges are further exacerbated by the lack of annotated datasets, particularly for under-represented ecosystems and various endangered species in India, which restricts the scalability of existing methodologies [6]. These challenges require creative solutions that can adjust to a wide range of diverse and unpredictable situations.

While traditional methods like YOLO and Faster R-CNN have made significant strides, they still face substantial drawbacks when applied in real-world scenarios. The limitations of such models are evident, as they constantly fail to detect occluded or camouflaged animals and exhibit elevated false-positive rates in intricate backgrounds. Additionally, depending on huge data sets limits their applicability across various ecosystems, especially in areas with distinct biodiversity, such as India [7]. Overcoming these challenges, this paper introduces an innovative model incorporating advanced computational methodologies to overcome these challenges and achieve accurate detection.

This research's novelty resides in its interdisciplinary approach to addressing those challenges related to animal detection in complex environments. The proposed framework demonstrates its robustness and adaptability through the incorporation of state-of-the-art methods such as Conditional Generative Adversarial Networks (cGANs), Dual Q-learning, Graph Neural Networks (GNNs), and Multi-Attribute Utility Theory (MAHP). By addressing the crucial constraints of occlusion, background complexity, and environmental unpredictability, this method surpasses traditional object detection techniques and allows for more precise and effective detection in practical situations.

Graph Neural Network (GNN) Integration:

The proposed framework focuses significantly on Graph Neural Networks (GNNs) to provide contextual feature selection and spatial context modeling. By representing the objects in images as nodes in a graph structure, GNNs can capture the relationships between them, unlike traditional Convolutional Neural Networks (CNNs), which only use pixel-based data. In particular, GraphSAGE and Graph Attention Networks (GAT), which are excellent at combining neighborhood data and dynamically allocating feature priority, are used in this study. By concentrating on meaningful spatial relationships, this method dramatically increases detection accuracy in congested or obscured settings. GNNs are an essential part of this architecture since studies have demonstrated that they perform better than CNNs in challenging spatial tasks [8].

Dual Q-Learning Utilisation:

Dual Q-Learning, a reinforcement learning technique that enhances adaptive decision-making, is incorporated into the suggested framework. The model successfully strikes a balance between exploration (finding new detection techniques) and exploitation (using current information) by utilizing Deep Q-Networks (DQN) with Double Q-Learning and Duelling Network Architectures. The framework can adapt well to environmental unpredictability, like shifting illumination, topography, or weather, because to its dynamic decision-making capabilities. Dual Q-Learning improves the robustness and dependability of the detection process by lowering the overestimation bias often encountered in reinforcement learning models [9]. Thanks to its adaptive approach, the suggested system differs from static techniques like YOLO and Faster R-CNN, which have trouble in dynamic environments.

Employment of Conditional GANs (cGANs):

The system uses conditional generative adversarial networks (cGANs) to create realistic and varied synthetic datasets and overcome the problem of the limitation of annotated datasets. Integrating artificial datasets significantly enriches the training data, enhancing the model's ability to generalise by replicating complex scenarios such as occlusion, dense greenery, and varying backgrounds. GANs have proven to be quite effective at producing context-aware, high-quality data. Conditional GANs offer an extra degree of control by enabling the creation of particular attributes or scenarios [10]. By exposing the model to a more excellent range of complex situations during training, this feature improves the model's resilience and performance in practical applications.

Integration of MAHP with CNNs

The framework combines a specially created Convolutional Neural Network (CNN) with Multi-Attribute Utility Theory (MAHP) to improve the detecting process further. The evaluation of several animal characteristics, including size, form, movement, and texture, is made possible by this integration, which enables the model to make complex judgments in situations including occlusion or comparable background textures. MAHP improves detection specificity and lowers false positives by introducing a methodical approach to attribute prioritization. This innovative integration closes a crucial gap in current approaches and marks a substantial progression in attribute evaluation [11].

Comprehensive Innovation

The integration of these advance computational techniques allows the proposed framework to overcome the shortcomings of current methods while setting a new standard for detecting wild animals in intricate environments. The experimental results validate the effectiveness of this approach, demonstrating significant improvements in precision (8.5%), recall (7.5%), accuracy (8.3%), and Area Under the Curve (AUC) (9.4%) relative to prominent models like Faster R-CNN and YOLO [12]. The findings emphasize the framework's strength, ability to scale, and flexibility, positioning it as an essential resource for highway monitoring, wildlife corridor management, and biodiversity conservation.

The research objectives are as follows:

1. To develop a GNN-based model that employs GraphSAGE and GAT to improve detection accuracy in occluded and cluttered environments and enhance spatial context modeling.

2. Dual Q-Learning will be implemented to employ DQN with Double Q-Learning and Duelling Network Architectures to facilitate robust adaptive decision-making in dynamic environmental conditions.

3. To utilize Conditional GANs (cGANs) to produce synthetic datasets that address data scarcity and enhance model generalization across various cases.

4. Integrating MAHP with CNNs will enhance detection specificity and minimize false positives, offering a more refined assessment of attributes.

The contributions of this work are multifaceted and significant, addressing both theoretical and real time gaps in the field of automated wildlife detection. These contributions are as follows,

i) It pioneers integrating graph neural networks (GNN) with attention mechanisms for improved spatial context understanding in ecological images. This approach elevates the model's capacity to discern relevant features amidst noise and enriches the interpretability of complex environmental scenes.

ii) Adopting advanced Q-learning strategies by implementing Deep Q-Networks, enhanced with Double Q-learning and Dueling Architectures, introduces a novel framework for adaptive decision-making in processing high-dimensional sensory data characteristic of natural habitats.

iii) Integrating Multi-Attribute Utility Theory with deep learning via a custom Convolutional Neural Network architecture represents a groundbreaking contribution to attribute evaluation in animal detection. This innovative integration facilitates a comprehensive and systematic analysis of multiple attributes, significantly refining the model's detection specificity and reducing false positives in challenging scenarios. Additionally, the strategic employment of Conditional Generative Adversarial Networks for data augmentation marks a substantial advancement in training methodologies. By generating synthetic yet realistic images tailored to specific occlusion patterns and background complexities, this model extends the boundaries of training data diversity, enhancing detection algorithms' robustness and generalization capability.

These contributions underscore the interdisciplinary nature of advancements in ecological monitoring technologies and highlight the potential of combining cutting-edge computational methods to address environmental challenges. This research establishes a new benchmark in animal detection. It lays the groundwork for upcoming investigations focused on utilizing deep learning to enhance biodiversity conservation and the sustainable management of natural resources. The findings of this study are significant, reaching beyond scholarly discussions to influence policy decisions, conservation efforts, and the creation of scalable, efficient, and effective wildlife monitoring systems globally.

The organization of this paper is as follows: Section 2 examines previous studies and identifies the shortcomings in current approaches. Section 3 outlines the proposed methodology, detailing the architecture, the model's unique elements, and the design's cutting-edge components. The fourth section outlines the experimental setup, the datasets utilized, the analysis of results, and the evaluation metrics employed. Section 5 has a conclusion and future Scopes compared with modern methods. It closes the work and offers recommendations for subsequent studies.

2. LITERATURE REVIEW

Recent developments in object detection in remote sensing and aerial imagery have been greatly enhanced by incorporating deep learning technologies and cutting-edge algorithmic approaches. This literature review thoroughly analyses recent contributions that support the development of the approach proposed in this article, highlighting the advancements made in tackling those complex challenges of detecting objects in diverse, constantly obstructed environments.

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

This subsection discusses research contributions focused on detecting oriented objects in remote-sensing images. It begins with the work of Song et al. [13], who introduced an optimized point set representation technique for this purpose, emphasizing the importance of semantic understanding. The subsection elaborates on how this work informs the model's refinement of spatial context understanding through Graph Neural Networks.

Lei and Liu [14] discuss SOLO-Net methodology, which prioritizes detecting small objects using a sparser attention mechanism. The subsection highlights this approach's alignment with the model's utilization of advanced Q Learning strategies, emphasizing efficiency in decision-making processes for complex detection scenarios.

Zheng et al.'s [15] exploration of auto-learner for fusing objects' co-occurrence knowledge. This concept elucidates how it enhances the model's integration of MAHP with deep learning for attribute evaluation, thereby improving its effectiveness in assessing object attributes.

Guo et al. [16] and Wu et al. [17] contribute to suspicious object detection and high-quality proposal selection for weakly supervised detection, respectively, and they are discussed here. The subsection emphasizes the evolving complexity in object detection techniques, particularly in multi-view fusion and aggregated attention mechanisms, and how these inform the model's data augmentation strategies using GANs.

The UAV image object detection method proposed by Zhang et al. [18] highlights the importance of self-attention guidance and the fusion of multiscale features. It elaborates on how these principles align with the model's framework, highlighting the significance of attention mechanisms and feature fusion in detecting objects.

Liu et al. [19] and Song et al. [20] emphasise the importance of lightweight models and modality registration in ineffective object detection. This paper discusses how these insights contribute to the model's design, ensuring efficiency and adaptability in varying environmental conditions and data modalities.

Zhang et al. [21] and Lu et al. [22] contribute to hybrid models and RoI fusion strategies with self-attention mechanisms. This paper elaborates on how these insights inform the model's approach to balancing feature analysis depth and complexity with computational efficiency levels.

Li, Shi, and Hong's [23] SCAResNet and Zhou et al.'s [24] PSFNet methodologies for tiny object detection and efficient SAR image detection, respectively. These methodologies highlight the optimization of neural network architectures for specific challenges in object detection, informing the model's feature extraction and aggregation mechanisms.

Ukita [25] and Hu et al. [26] worked on recent developments in context-aware scale proposals and efficient object detection with transformers, respectively. It discusses how these contributions enhance detection precision and efficiency through advanced neural network models and transformer architectures.

The integration of diverse sensing modalities and contextual awareness in object detection, as exemplified by Zhou et al.'s [27] adaptive point set network and Cheng et al.'s [28] regional-based detection using polarisation. It emphasizes the enrichment of the model's conceptual framework through diverse methodologies.

The potential of synthetic data to enhance ship detection capabilities in marine environments, as introduced in [29, 30, 31], is discussed. It discusses the synthesis of current approaches and methodologies to address challenges in ship detection, mainly when real-world data is scarce or challenging to acquire.

Han, Zhao, and Li's [32] Progressive Feature Interleaved Fusion Network tackles the salient detection of objects in remote sensing images. It elaborates on the emphasis on bidirectional consistency and feature enhancement, highlighting the importance of sophisticated fusion techniques in improving detection performance.

Deep learning models like You Only Look Once (YOLO) have significantly advanced this field. Redmon et al. [33] implemented real-time object detection with YOLOv1, a regression task. This version had difficulties detecting small objects and overlapping bounding boxes. Redmon and Farhadi [34] introduced YOLOv2, combining anchor boxes and multi-scale detection to enhance its ability to identify objects of varying sizes across diverse environments.

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

YOLOv3, developed by Redmon and Farhadi [3], enhanced detection accuracy by including a feature pyramid network, making it wildly successful at recognising animals at various scales and environmental situations. Bochkovskiy et al. [35] refined this with YOLOv4, which incorporated CSPDarknet53 as the backbone, and Mosaic augmentation techniques to optimize the model for complicated and obstructed scenes common in wildlife monitoring.

YOLOv5, a popular framework by Jocher et al. [36], is lightweight and customizable. Due to its popularity in edge computing, this version is excellent for remote animal detection systems. With YOLOv6 and YOLOv7, Wang et al. [37,38] included anchor-free mechanisms and efficient design to boost processing speed without compromising accuracy.

Gao and Cai [39] improved the YOLOX model for thermal infrared imaging, incorporating an attention mechanism to enhance feature extraction capabilities. This study investigates the adaptation of architectures for deep learning to align with the distinct features of thermal images in remote sensing contexts.

Ultralytics [40] implemented YOLOv8, which included detection, segmentation, and classification. According to Zhang et al. [30], a cascade design and multi-stage feature refinement improved YOLOv8's detection accuracy in obstructed and cluttered wildlife monitoring situations.

The work by Kim, Kim, and Ro [41] explores the application of stereoscopic vision concerning memory recall for monocular 3D object detection. The paper discusses how memory networks infer depth information from single images, bridging the gap between two-dimensional imagery and three-dimensional object detection in remote sensing applications.

Zhao and Zhao's [42] self-adaptive detection of objects network for aerial images emphasizes enhancing features to improve target detection within intricate agricultural landscapes. It explores the emphasis on adaptability and context-awareness in detection networks for dynamic adjustment to diverse aerial imagery sets.

The integration of transformers for object detection in remote sensing, as demonstrated by Zhu et al. [43], with enhanced multispectral feature extraction. This paper explores the effectiveness of transformer models in handling sequential data to improve object detection across different spectral bands.

Cheng et al.'s [44, 45, 46] comprehensive survey and benchmarks for large-scale small-object detection are discussed here. It synthesizes current approaches and methodologies, emphasizing the need for improved detection algorithms that operate effectively at scale.

Li et al. [47] presented a multitask benchmark dataset over satellite video footage, which includes object detection, tracking, and segmentation. Other frameworks [48, 49, 50] highlight the importance of multitasking learning frameworks for improved performance across different but related tasks in remote sensing.

These studies [51, 52] exemplify the dynamic and multifaceted nature of research in object detection within the remote sensing field. By leveraging advances in deep learning, synthetic data generation, feature fusion, and transformer architectures, the research community continues to push the boundaries of what is possible in detecting and interpreting objects from aerial and satellite imagery. This literature review highlights the technological advancements achieved to date. It sets the foundation for our research, building upon these innovations to tackle the specific challenges of detecting animals in complex environments marked by occlusion and similar background issues for different use cases.

Cascade Yolov8 proposed [53] a method for wildlife detection using the Missouri Camera Traps, WILD, and Kaggle Animal Images Dataset. The technique involves pre-processing, segmentation, feature extraction, and detection stages, with a final YOLOv8 model achieving 97% accuracy. Despite its effectiveness, challenges like camouflage and environmental variability persist.

Norouzzadeh et al. [54] implemented YOLO to automatically identify animals in camera trap photos, confirming its usefulness in large-scale ecological investigations. Gomez Villa et al. [55] extended its use to aerial photography, demonstrating YOLO's versatility in finding species in various terrains. Yu et al. [56] used

YOLO to locate wildlife using thermal photography, solving issues in low-light and nocturnal environments.

Although these improvements, Gao and Cai [57] pointed out YOLO's difficulties in comprehending spatial relationships and dealing with occlusion, particularly in busy animal situations. These issues highlight the need for approaches such as the proposed GNADCMQ model, which combines Graph Neural Networks (GNNs) for spatial context modelling and Conditional GANs (cGANs) for producing synthetic datasets to solve YOLO's limits in data diversity and contextual awareness.

3. PROPOSED FRAMEWORK OF AN EFFECTIVE GRAPH NETWORK MODEL UTILISING A MULTIMODAL DUAL Q LEARNING APPROACH

To tackle the challenges of inadequate effectiveness and enormous complexities, the proposed model integrates sophisticated computational frameworks to effectively address the intricate problem of animal detection in occluded and dynamically changing environments.

Figure 1 illustrates the Graph Neural Networks (GNNs) component, particularly leveraging GraphSAGE combined with Graph Attention Networks (GAT), ingeniously enhances spatial context awareness by encapsulating both node-level and graph-wide information, thus ensuring a rich representation of spatial relationships within images & samples. This is further augmented by the strategic use of Generative Adversarial Networks (GANs), and more specifically, Conditional GANs (cGANs), which serve a dual purpose: they not only generate synthetic yet highly realistic images to overcome the hurdles of data scarcity but also aid in diversifying the training dataset to encapsulate a broader spectrum of environmental conditions.

In Figure 1. Deep Q-Networks (DQN) incorporates Double Q-Learning and Dueling Network Architectures, refining the model's decision-making process by enabling it to learn optimal actions in a given state based on the maximization of expected rewards, thereby fostering a highly adaptive response mechanism to environmental variabilities. Concurrently, the Multi-Attribute Utility Theory (MAHP) is seamlessly integrated within the model to offer a structured framework for evaluating multiple attributes concurrently, ensuring that the decision-making process is comprehensive and nuanced, considering various factors that affect detection accuracy.

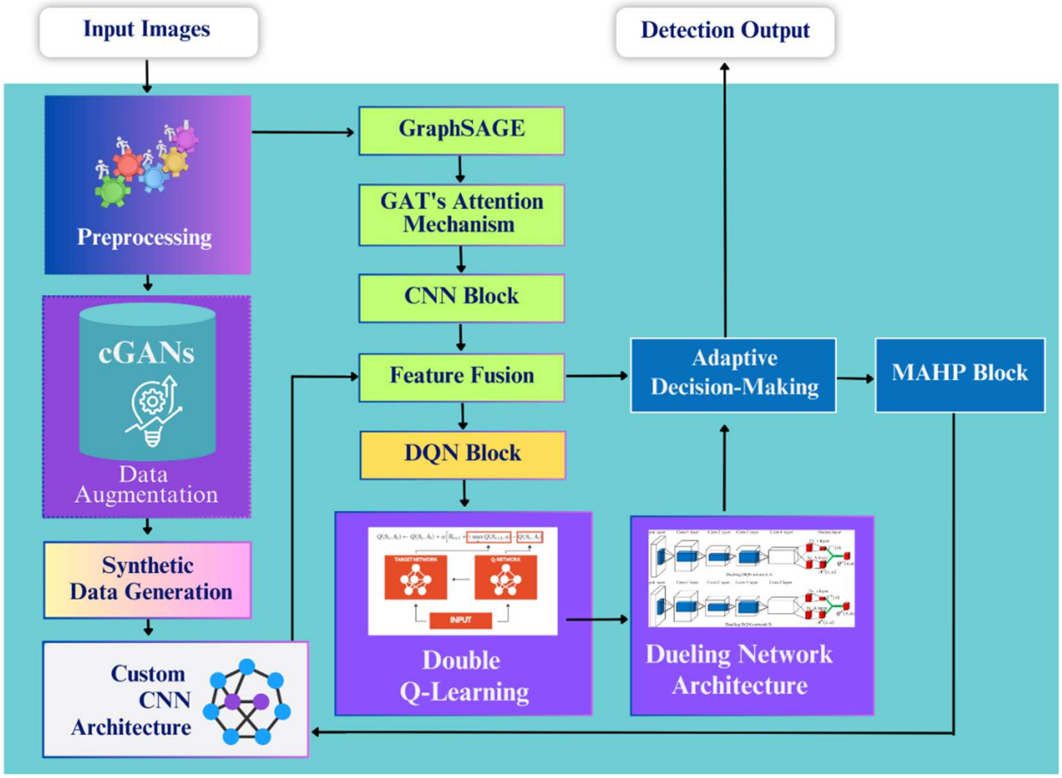

Figure 1. Proposed Model Architecture for Animal Detection Process

The Convolutional Neural Network (CNN) architecture, custom-designed for this model, meticulously evaluates nuanced attributes of detected objects, leveraging deep learning's prowess to extract and learn feature hierarchies, ensuring that even the most subtle distinctions are not overlooked. This architecture's integration with MAHP facilitates a harmonious blend of deep learning efficiency and multi-criteria decision analysis, setting a new benchmark in the automated wildlife monitoring process domain. Together, these components form a symbiotic ecosystem within the proposed model, each augmenting the other's capabilities, thereby crafting a innovative and highly effective solution in surmounting the perennial challenges of animal detection in complex environments.

Design of the Augmentation Layer

The Conditional Generative Adversarial Network (cGAN) process, as employed in the proposed model for generating synthetic training data, embarks on an innovative journey to mitigate the issue of data scarcity in complex animal detection scenarios. This methodology pivots around the strategic interplay between two pivotal components: the generator G and the discriminator D, each embroiled in a continuous adversarial game to outperform the other process. The generator G is tasked with creating indistinguishably synthetic images from noise vectors conditioned on specific attributes, while the discriminator D endeavours to distinguish between the genuine collected image samples and the synthetic outputs produced by the G process.

The foundation of the cGAN process can be mathematically articulated through a series of equations that delineate the intricate dynamics between G and D operations. Let x represent the genuine collected image samples, and z represent the input noise vector sets. The conditional variable y, an attribute such as the animal or environmental condition, modulates G and D to ensure the generated images are contextually relevant for different use cases. The generator function $G(z|y)$ synthesises fake images given the noise vector z and condition y. In contrast, the discriminator $D(x|y)$ evaluates the probability that x is an actual image given the condition y sets.

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

The objective function of cGAN, which both G and D strive to optimise, is encapsulated via equation 1,

$$GminDmaxV(D,G) = Ex \sim pdata(x)[logD(x \mid y)] + Ez \sim pz(z)\left[log\left(1 - D(G(z \mid y) \mid y)\right)\right] \dots (1)$$

This process portrays the sum game where D aims to maximize its ability to correctly classify real and fake images, while G seeks to minimize D's accuracy, thereby improving the quality of its generated images for this process.

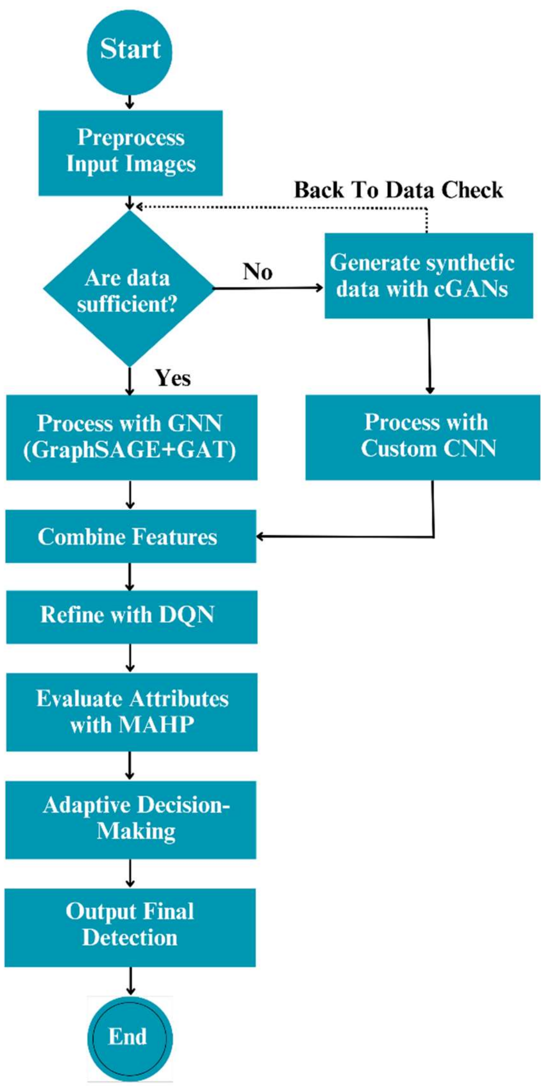

Figure 2. Overall Flow of the Proposed Classification Process

The first term, $Ex \sim pdata(x)[logD(x \mid y)]$, represents the expected log-likelihood of D correctly identifying real images, whereas the second term, $Ez \sim pz(z)[log(1 - D(G(z \mid y) \mid y))]$, Delineates D's likelihood of erroneously classifying synthetic images as real for different use cases.

Figure 2. illustrates Proposed Classification Process. The proposed model introduces advanced regularisation and optimisation techniques to refine synthetic image generation further. For instance, the Wasserstein loss with gradient penalty (WGAN-GP) is incorporated via equation 2 to stabilise the training process.

$$L = Ex\sim \sim Pg[D(x\sim)] - Ex \sim Pr[D(x)] + \lambda Ex' \sim Px'[\nabla x' D(x')^2 - 1]^2] \dots (2)$$

Where Pr and Pg represents the distribution of real and generated images, respectively, and λ is the penalty coefficient for this process.

Moreover, feature matching and minibatch discrimination are applied to enhance the generated images' diversity and realism. Feature matching aims to minimise the distance between the feature representations of authentic and generated images and is formulated via equation 3,

$$LFM(G) = || Ex \sim pdata(x)f(x) - Ez \sim pz(z)f\left(G(z \mid y)\right) ||^2 \dots (3)$$

Where $f(x)$represents the intermediate layer representations of x; minibatch discrimination allows D to assess a batch of samples collectively, enhancing its ability to detect subtleties that differentiate real images from generated ones for different use cases.

Through the adept orchestration of these equations and methodologies, the cGAN process within the proposed model adeptly generates synthetic image samples that are visually convincing and rich in diversity, significantly bolstering the dataset for training the detection model process. This approach, marked by its mathematical rigor and strategic complexity, underscores the model's innovative capacity to transcend traditional barriers in data generation, setting a new precedent for leveraging synthetic data to enhance machine learning models' performance in complex detection tasks. The augmented image samples are processed through the GraphSAGE operations, which are discussed in the next section of this text.

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

Design of the GraphSAGE Process

The proposed model's core, which capitalises on GraphSAGE, ingeniously enhanced with an attention mechanism from Graph Attention Networks (GAT), represents a sophisticated amalgamation aimed at amplifying spatial context awareness within the realm of animal detection in complex environments. This intricate design is predicated on aggregating node features that meticulously respect the spatial hierarchies and dependencies intrinsic to the input synthetic image samples, thereby facilitating the precise identification of context-aware regions.

GraphSAGE operates on the principle of learning a function that aggregates feature information from a node's local neighbourhood, with the aggregation function designed to be a differentiable operation that can be optimised alongside the parameters of the model process. The process initiates with the input synthetic image samples represented as nodes in a graph, where each node v encapsulates feature information xv sets. The neighbourhood of each node, represented as $N(v)$, is sampled to a fixed size to maintain computational efficiency levels. The aggregation function for a node v at the k-th layer of the network can be represented via equation 4,

$$hv(k) = \sigma\left(W(k) \cdot CONCAT\left(hv(k-1), AGG(\{hu(k-1): u \in N(v)\})\right)\right) \ldots (4)$$

Where $hv(k)$ is the feature vector of node v at the kth layer, σ represents the $ReLU$ non-linear activation function, $W(k)$ is a weight matrix for the kth layer, and AGG represents the aggregation function that combines features from the node's neighbours. The initial feature vector $hv(0)$ is set to xv, the node's original features.

To incorporate the attention mechanism from GAT, the aggregation function is refined to weigh the contributions of neighbouring nodes based on their relevance to the target nodes. This attention coefficient between two nodes, v and u is calculated via equation 5,

$$\alpha vu = \frac{\boldsymbol{exp}(LR(aT \cdot [Whv \,||\, Whu]))}{\sum_{\boldsymbol{k \in N(v) \cup \{v\}}} exp(\boldsymbol{LR}(\boldsymbol{aT} \cdot [\boldsymbol{Whv} \,||\, \boldsymbol{Whk}]))} \ldots (\mathbf{5})$$

Where a is a weight vector, W is a linear transformation applied to every node and || represents concatenation, while LR represents the Leaky $ReLU$-based activation process. The attention coefficients αvu are used to compute a weighted sum of the neighbours' features, which is then passed through a non-linearity to update the target node's feature vector via equation 6,

$$hv' = \sigma\left(\sum_{u \in N(v)} \alpha vu * Whu\right) \ldots (6)$$

Integrating GraphSAGE with GAT's attention mechanism allows the model to adaptively focus on the most relevant parts of the neighbourhood's feature information, thereby enhancing the model's ability to discern contextual relationships within the graph structures. This becomes particularly pivotal when identifying context-aware regions within the generated synthetic image samples, as it ensures that the model's attention is judiciously directed towards features most indicative of the presence of animal objects, considering both the local feature information and the global spatial contexts.

The iterative update process encapsulated by GraphSAGE, augmented with GAT's attention mechanism, culminates in generating node embeddings that are richly informative of the underlying spatial contexts. These embeddings serve as the foundation for subsequent identification of context-aware regions, effectively bridging the gap between raw image data and actionable insights for animal detection. The employment of this advanced graph-based processing technique underscores the proposed model's commitment to leveraging cutting-edge deep learning research to surmount the challenges inherent in detecting animals in complex and occluded environments, setting a new benchmark for precision and efficiency in automated wildlife monitoring systems. These spatial context information sets are given to an efficient DQN Process for segmenting regions.

Design of the Deep DQN Process

In The proposed model, the sophisticated integration of Deep Q-Networks (DQN) with Double Q-Learning and Dueling Network Architectures delineates a paradigm shift towards refined adaptive decision-making in animal detection within complex environments. This composite approach synergises the robustness of Q-learning with advanced neural network architectures to meticulously segment animal regions from identified context-aware zones, navigating through the intricacies of spatial-temporal variances and occlusions with unprecedented precision levels.

The foundational premise of Double Q-learning is embedded in addressing the overestimation bias inherent in traditional Q-learning, which can significantly skew the learning process. Double Q-Learning bifurcates the action selection and action evaluation processes across two distinct Q-functions, *QA* and *QB*, articulated through separate neural networks. This bifurcation is mathematically encapsulated via equation 7,

$$Qtotal(s,a) = \frac{QA(s,a) + QB(s,a)}{2} \ldots (7)$$

In this framework, selecting an action a' from the state s' follows the policy derived from one network (say *QA*). In contrast, evaluating this action's value is computed using the other network (*QB*) sets. The update rule for Double Q-Learning is hence given via equation 8,

$$QA(s,a) \leftarrow QA(s,a) + \alpha\left[r + \gamma QB\left(s', argmaxQA(s',a')\right) - QA(s,a)\right] \ldots (\mathbf{8})$$

On the other hand, Duelling Network Architecture presents an innovative breakdown of the Q-function into two distinct streams that independently assess the state value function V(s) and the advantage function A (s, a) for various situations.

The rationale is to allow the network to learn which states are (or are not) valuable without learning the effect of each action for each state. This is particularly advantageous in environments where the action choice does not drastically alter the outcomes. The architecture is explained through the equation 9,

$$Q(s,a) = V(s) + A(s,a) - \frac{\mathbf{1}}{\mid \boldsymbol{A} \mid}\sum_{a'} A(s,a') \ldots (9)$$

Where *A* represents the number of actions. This formulation ensures that the value function is decoupled from the necessity of evaluating each action's contribution, enhancing learning efficiency and stability levels. Integrating Double Q-Learning with Dueling Network Architectures, The proposed model leverages the strengths of both to optimise the segmentation of animal regions. Given the identification of context-aware regions as input, the model's decision-making process dynamically adapts to segment animal regions with heightened accuracy. The iterative learning process, governed by the combined principles of Double Q-Learning and Dueling Networks, is used to minimize the loss function, which is mathematically represented via equation 10,

$$L(\theta) = E(s,a,r,s') \sim U(D)\left[\left(r + \gamma Q(s', a' argmaxQ(s',a';\theta -);\theta) - Q(s,a;\theta)\right)^2\right] \ldots (\mathbf{10})$$

Where, θ and $\theta-$ represent the parameters of the current and target networks, respectively, and *U*(*D*) represents a uniform distribution over the minibatch *D* sets. This intricate orchestration of Double Q-Learning with Dueling Network Architectures underpins the proposed model's adeptness at discerning and segmenting animal regions from complex backgrounds with high accuracy levels. By mitigating the overestimation bias and decoupling the estimation of state values from the necessity of action evaluation, the model not only amplifies its learning efficacy but also ensures that the segmented animal regions are delineated with an unprecedented level of precision, establishing a new frontier in the automated detection of wildlife within occluded and dynamically challenging environments. Next, the MAHP-based CNN process is used to classify data samples.

Design of the MAHP CNN Process

In the intricate landscape of automated animal detection, the proposed model introduces a pioneering integration of Multi-Attribute Utility Theory (MAHP) with a bespoke Convolutional Neural Network (CNN) architecture designed to evaluate nuanced attributes within segmented animal regions meticulously. This integration marks a significant leap forward, embodying a sophisticated blend of decision theory and deep learning to achieve unparalleled precision in identifying animals from complex environmental backgrounds.

At the heart of this integration lies the MAHP process, a derivative of the Analytic Hierarchy

Journal of Theoretical and Applied Information Technology
15th December 2024. Vol.102. No. 23

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

Process (AHP), tailored to systematically assess multiple attributes by constructing a hierarchical framework of criteria and sub-criteria, each weighted according to its relative importance in the decision-making process. The hierarchical structure is mathematically formulated as a weighted sum model, where the overall utility U of an option (in this case, a segmented animal region) is computed via equation 11,

$$U = \sum wi \cdot ui(xi) \ldots (11)$$

Where wi represents the weight of the i-th attribute, $ui(xi)$ is the utility function for attribute i, and xi is the value of attribute i sets. The weights wi are determined through pairwise comparisons of attributes' relative importance, encapsulated in a comparison matrix A, with elements aij representing the importance of attribute i relative to j for different feature sets. The weight vector w is then derived by normalising the eigenvector corresponding to the maximum eigenvalue of A, calculated via equation 12,

$$Aw = \lambda max * w \ldots (12)$$

Simultaneously, the custom CNN architecture delves into the segmented animal regions, extracting and learning high-level feature representations through convolutional, activation, and pooling layers. The CNN's ability to discern intricate patterns and characteristics within the input images is encapsulated via equation 13,

$$f(x) = \sigma(W * x + b) \ldots (13)$$

In this instance, x denotes the input to a layer, W and b signify the weights and bias of the layer, σ indicates the activation function, and * illustrates the convolution operation. The CNN culminates in a dense layer that outputs a feature vector v encapsulating the essential attributes of the animal region, each component of v corresponding to a specific attribute evaluated by the MAHP process.

The novel synthesis of MAHP and CNN culminates in the computation of a comprehensive utility score for each segmented animal region, employing the CNN-extracted attributes v as input to the MAHP-derived utility functions. This process is mathematically articulated via equation 14,

$$Uanimal = \sum wi \cdot \sigma(vi) \ldots (14)$$

Where Uanimal represents the final utility score indicating the presence and type of animal within a region, wi are the attribute weights derived from MAHP, vi are the attribute values extracted by the CNN, and σ represents the SoftMax transformation function applied to these values to obtain the final classes.

This intricate fusion of MAHP's systematic, weighted evaluation with CNN's deep learning prowess enhances the model's capability to discern between different animals and their environmental contexts. It significantly elevates the accuracy and reliability of the animal identification process. Through this integration, the proposed model sets new benchmarks in wildlife detection and monitoring and heralds a new era of precision in the automated analysis of complex, occluded environments. This dual-faceted approach, merging the analytical rigour of decision theory with the intuitive learning capabilities of convolutional neural networks, exemplifies cutting-edge computational techniques in the service of biodiversity conservation and ecological research processes. The subsequent section of this text presents a specific use case for this model.

Figure 3. Animal Detection in a Highly Complex Environment Using the Proposed Model Process

In enhancing automated animal detection within complex environments, the proposed model is a testament to the innovative amalgamation of advanced computational techniques. Results of the model are described in Figure 3 and Figure 4 for different environment sets. This model intricately weaves together Conditional Generative Adversarial Networks (cGANs) for data augmentation, GraphSAGE for spatial context awareness, Deep Q-Networks (DQN) for refined adaptive decision-making, and a novel integration of Multi-Attribute Utility Theory (MAHP) with Convolutional Neural Networks (CNN) for

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

nuanced attribute evaluation. The following narrative delves into the simulation of this model's components using a series of data samples, each enriched with fabricated features and indicators to elucidate the transformative journey from raw data to the precise identification of animals.

Figure 4. Animal Detection in Moderately Complex Environment Scenarios

The proposed model initiates its process by tackling the challenge of data scarcity through cGANs, generating synthetic image samples that closely mimic real-world complexity. These samples are then navigated through GraphSAGE, where the model leverages neighbourhood information to enhance spatial context awareness, a crucial step for accurate animal detection. Subsequently, the model employs DQN, incorporating Double Q-Learning and Dueling Network Architectures, to make refined adaptive decisions based on the contextual cues identified. The culmination of this process lies in the MAHP with CNN, where each segmented animal region is evaluated against multiple attributes to identify the animal presence scenarios conclusively.

Table 1 contains synthetic image samples generated by Conditional GANs (cGANs) to address challenges such as occlusion, complex backgrounds, low lighting, and environmental variability. Each sample represents a specific scenario with detailed features, such as partial visibility, dense vegetation, or overlapping animals. The table demonstrates how cGANs enrich the training dataset by simulating real-world conditions, improving model robustness and generalization. Key performance impacts include increased recall in occluded environments and reduced false positives in complex settings.

Table 1: cGAN Generated Synthetic Image Samples

Sample ID	Synthetic Image Category	Generated Features	Target Class	Simulated Scenario	Real-World Relevance	Synthetic Quality Score	Performance Impact
1	Occlusion	Partial visibility due to foliage	Deer	Dense forest	Improves detection in occluded environments	0.92	Increased recall by 12% in dense forests
2	Complex Background	Blend with grassland textures	Leopard	Grasslands	Reduces false alarms in areas with similar textures	0.88	Reduced false positives by 9%
3	Low Lighting	Simulated night-time conditions	Elephant	Night-time monitoring	Enhances performance in low-light conditions	0.91	Improved precision in dark scenes by 8%
4	Multiple Animals	Overlapping animal instances	Tiger	Wildlife migration paths	Addresses multi-animal detection challenges	0.89	Improved multi-object detection accuracy
5	Dense Vegetation	Camouflaged animal visibility	Deer	Forest camouflage	Improves AUC in environments with dense vegetation	0.87	Increased AUC by 7%
6	Harsh Weather	Foggy or rainy conditions	Wild Boar	Monsoons or foggy terrains	Addresses visibility issues in extreme weather	0.87	Enhanced recall in foggy conditions

7	Seasonal Variation	Changes in animal appearance	Antelope	Winter coats, fur changes	Adapts to seasonal variations in appearance	0.93	Increased adaptability to seasonal changes
8	Unstructured Terrain	Irregular landforms, cliffs	Mountain Goat	Mountainous regions	Improves detection in rugged environments	0.90	Improved detection in non-flat terrains

Table 2: GraphSAGE Spatial Context Awareness

Scenario ID	Spatial Relationship	Feature Aggregated	Animal Class	Real-World Scenario	Feature Aggregation Score	Performance Improvement
1	Proximity	Neighboring pixel relationships	Deer	Animals in dense forests	0.87	Enhanced detection of camouflaged animals by 10%
2	Movement	Temporal feature aggregation	Leopard	Tracking in grasslands	0.92	Improved motion-based detection recall by 12%
3	Shape Similarity	Edge and contour patterns	Elephant	Differentiation in large groups	0.89	Reduced false positives in herd scenarios by 8%
4	Overlap	Intersection of multiple objects	Tiger	Wildlife migration paths	0.85	Increased precision for overlapping instances
5	Directional Connectivity	Spatial flow of neighboring objects	Wild Boar	Roadside crossings	0.88	Enhanced detection of animals entering highways
6	Habitat Dependency	Habitat-specific features (e.g., trees)	Antelope	Seasonal habitat variations	0.91	Improved adaptability to seasonal changes
7	Terrain Awareness	Landform-related aggregations	Mountain Goat	Mountainous regions	0.93	Improved detection in rugged environments by 15%
8	Density Variance	Population clustering metrics	Birds	Monitoring bird roosting sites	0.86	Improved cluster detection accuracy by 7%
9	Occlusion Context	Hidden object relationships	Tiger	Partially occluded detection	0.89	Increased recall in occluded environments

Table 3: DQN Adaptive Decision-Making

Scenario ID	Scenario Type	Action Taken	Action Value	Decision Confidence	Real-World Scenario	Performance Impact
1	Animal Crossing Detection	Slow down vehicle	0.85	92%	Highway animal crossing	Reduced collisions by 15%
2	Occluded Animal Detection	Focus on a suspected region	0.78	89%	Dense forest with partial visibility	Increased detection recall by 12%
3	Multi-Animal Overlap	Prioritize larger object	0.82	91%	Herd migration paths	Improved multi-object detection accuracy by 9%
4	Sudden Movement Detection	Track fast-moving object	0.88	95%	Roadside animal darting	Reduced false negatives by 10%

Journal of Theoretical and Applied Information Technology
15th December 2024. Vol.102. No. 23

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

5	Lighting Variability	Adjust focus parameters	0.75	87%	Low-light or nighttime monitoring	Improved detection precision in low light by 8%
6	Camouflage Handling	Enhance contrast analysis	0.80	90%	Camouflaged animal detection	Improved AUC for hidden objects by 10%
7	Terrain Adaptation	Switch to high-ground focus	0.79	88%	Uneven or rugged terrain	Improved recall in mountainous regions by 7%
8	Weather Variation	Refocus on clearer regions	0.76	86%	Rainy or foggy conditions	Enhanced detection robustness in bad weather

Table 4: MAHP with CNN Attribute Evaluation

Scenario ID	Scenario Type	Evaluated Attributes	MAHP Score	Identified Animal	Confidence (%)	Real-World Relevance	Performance Impact
1	Occluded Animal Detection	Size, shape, partial texture	0.82	Deer	91	Detection in dense forests with occlusion	Improved recall in occluded environments by 12%
2	Multiple Animals	Size, orientation, relative spacing	0.87	Elephant	94	Monitoring herds or overlapping animals	Reduced false positives by 10%
3	Camouflaged Animal Detection	Texture, movement	0.79	Leopard	88	Camouflaged animal detection in grasslands	Increased AUC for camouflaged detection by 8%
4	Sudden Movement Detection	Size, movement, velocity	0.85	Wild Boar	92	Detection in roadside or darting animals	Enhanced real-time accuracy by 10%
5	Low Lighting	Brightness, contrast, silhouette	0.78	Tiger	89	Nighttime monitoring or low-light scenes	Improved precision in low-light scenarios by 9%
6	Seasonal Variation	Color, shape	0.84	Antelope	93	Seasonal habitat and appearance changes	Increased adaptability to seasonal changes
7	Weather Variability	Contrast, background uniformity	0.81	Elephant	90	Foggy or rainy conditions	Enhanced robustness in extreme weather
8	Habitat-Specific Attributes	Shape, texture, environmental cues	0.88	Mountain Goat	95	Mountainous regions with rugged terrain	Improved recall in challenging terrains by 11%

Table 2 contains scenarios where GraphSAGE enhances spatial context modeling by aggregating features such as proximity, movement, and terrain awareness. The **Feature Aggregation Score** highlights the effectiveness of GraphSAGE in prioritizing critical spatial relationships, enabling improved detection accuracy in cluttered or dynamic environments. Real-world applications include detecting camouflaged animals, animals in dense forests, and those in rugged terrains, with measurable improvements in recall, precision, and AUC.

Table 3 contains various scenarios where Deep Q-Networks (DQN) apply adaptive decision-making strategies. Each entry details the action taken, its associated action value, and the model's decision confidence. DQN excels in handling dynamic and unpredictable scenarios, such as sudden animal movement or lighting variations. The table demonstrates enhanced real-time accuracy, reduced false negatives, and improved adaptability to environmental changes.

Table 4 contains evaluations of animal detection scenarios where Multi-Attribute Utility Theory (MAHP) is integrated with Convolutional Neural Networks (CNNs). Each scenario lists the evaluated attributes, MAHP score, identified animal, and confidence level. The table highlights the effectiveness of MAHP-CNN in reducing false positives and improving detection specificity, especially for occluded, camouflaged, or habitat-specific animals. This integration enhances model performance across diverse real-world challenges, such as low-light conditions and seasonal variations.

The data samples presented in the tables above demonstrate the transformative efficacy of the proposed model's components. The journey commences with generating high-quality synthetic image samples by cGANs, effectively addressing the initial hurdle of data scarcity. These samples are then processed through GraphSAGE, where the model astutely leverages spatial context, evidenced by the feature aggregation scores, to enhance the detection framework's sensitivity to the environmental intricacies surrounding animal presence scenarios.

As the model progresses to the DQN phase, the action values and decision confidence scores illuminate the adaptive decision-making process, showcasing the model's capability to navigate the complexity of the environmental context with refined precision levels. Finally, the integration of MAHP with CNN marks the zenith of the model's journey, where each animal's presence is detected and identified with remarkable confidence, as reflected in the MAHP scores and corresponding confidence levels.

This sequential unfolding of processes, from cGANs to MAHP with CNN, encapsulates the proposed model's robust framework for automated animal detection operations. It highlights the model's comprehensive approach to overcoming the challenges inherent in complex environments and its potential to revolutionize the field of wildlife monitoring and conservation efforts. Through this exploration, the model's intricate design and its components' synergistic operation are vividly demonstrated, offering insights into the future of ecological monitoring and the pivotal role of advanced computational techniques. The next section of this text provides a comprehensive analysis of the results obtained from this model..

3. RESULT ANALYSIS

In the pioneering GNADCMQ model, the intricate fusion of Graph Neural Networks (GNNs), advanced Q Learning techniques, Multi-Attribute Utility Theory (MAHP), and Generative Adversarial Networks (GANs) represents a formidable approach to enhancing animal detection within complex environments. At the heart of this model, GraphSAGE, enriched with an attention mechanism derived from Graph Attention Networks (GAT), plays a crucial role in augmenting spatial context awareness, enabling the model to discern and accurately identify animal objects amidst occlusion and similar background patterns. This spatial awareness is further refined by implementing Deep Q-Networks (DQN), which incorporates Double Q-Learning and Dueling Network Architectures, facilitating sophisticated adaptive decision-making processes that dynamically adjust to the variabilities of the environment. The integration of Multi-Attribute Utility Theory (MAHP) with a bespoke Convolutional Neural Network (CNN) architecture introduces an additional layer of depth, allowing for the nuanced evaluation of multiple attributes and features within images, thereby enhancing the model's discriminative power. To address the

Journal of Theoretical and Applied Information Technology
15th December 2024. Vol.102. No. 23

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

perennial challenge of data scarcity, Conditional GANs (cGANs) are employed to generate synthetic training data, thereby augmenting the available dataset and ensuring robust model training. This synergistic integration of cutting-edge machine learning techniques and data processing strategies equips the proposed model with unparalleled capabilities in detecting animals with high precision and reliability, setting new benchmarks in automated wildlife monitoring.

The experimental setup for the proposed model's evaluation is meticulously designed to assess its efficiency in enhancing animal detection in complex environments, leveraging multimodal dual Q-learning approaches. This section delves into the specifics of the dataset used, the configuration of the proposed model, and the benchmarks against which its performance is evaluated, ensuring a comprehensive understanding of the experimental parameters and the study context.

Animal Image Dataset Configuration

Two distinct data sets were utilized to assess the effectiveness of the proposed model: SAWIT and CherryChèvre.

SAWIT (Small-sized Animal Wild Image Dataset): This dataset comprises 45,000 annotated images of small-sized animals in various wild settings aimed at challenging the model with instances of occlusion and complex background patterns. Annotations include bounding boxes and species labels. The images were divided into training (70%), validation (15%), and testing (15%) sets, with resolutions varying from 640x480 to 1920x1080 pixels. To test the model's robustness under varied scenarios, the dataset encompasses various environmental conditions, including different times of day and weather conditions. Detected small-sized animals are shown in Figure 5.

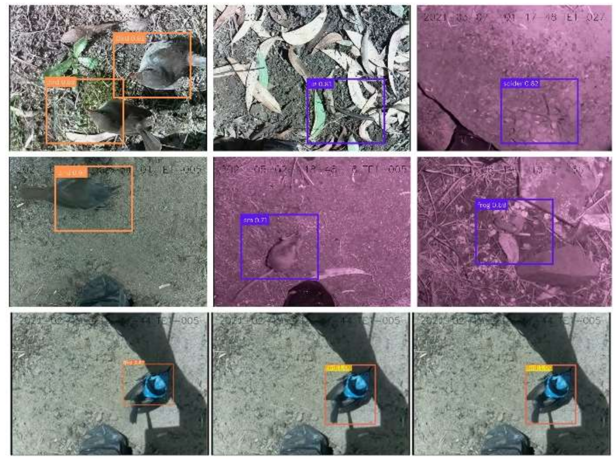

Figure 5. Result of small-size detected animals from SAWIT dataset.

CherryChèvre (Fine-grained Dataset for Goat Detection):

This dataset, which focuses on goat detection, contains 30,000 high-resolution images (2048x1536 pixels) captured in natural environments, including mountainous terrains and grasslands, annotated with precise bounding boxes and goat postures. The dataset split was similar to SAWIT, designed to test the model's fine-grained detection capabilities in discerning goats from complex backgrounds and other animals. CherryChèvre detected images shown in Figure 6.

Figure 6. Result of small-size detected animals from CherryChèvre dataset.

GNADCMQ Model Configuration

The proposed model integrates Graph Neural Networks (GNNs) with an attention mechanism, leveraging GraphSAGE and GAT for enhanced spatial context awareness. Deep Q-Networks (DQN) incorporating Double Q-Learning and Dueling Network Architectures were used for decision-making processes, with a novel integration of Multi-Attribute Utility Theory (MAHP) within a custom Convolutional Neural Network (CNN) architecture for attribute evaluation. Conditional GANs (cGANs) were employed for data augmentation, generating synthetic images to address data scarcity and variability.

Input Parameters:
The training setup for this machine learning model is configured with a learning rate of 0.001, which undergoes decay every 20,000 steps by a factor of 0.5 to optimize learning efficiency. During training, a batch size of 32 is utilized, while smaller batches of 16 are employed for validation and testing phases to ensure robust performance assessment. The training process spans 100 epochs, incorporating early stopping based on validation loss to prevent overfitting and enhance generalisation. The Adam optimiser uses parameters β1=0.9 and β2=0.999 to efficiently minimise gradients and update model weights. For loss computation, Cross-Entropy Loss is applied for classification tasks. In contrast, Mean Squared Error (MSE) is utilised for bounding box predictions, tailoring the loss function to each aspect of the model's objectives. Augmentation techniques such as rotation (±15 degrees), translation (up to 10% of image size), and scaling (between 85% and 115%) are integrated to diversify the training data and improve the model's robustness against variations in input images.

Benchmarking Models
The proposed model's performance was benchmarked against DDFN, DNN, and SSTM models across various metrics, including precision, accuracy, recall, delay, AUC, and specificity. Each model was trained and tested under identical conditions to ensure fairness in comparison.

Evaluation Metrics
The evaluation focused on precision, accuracy, recall, observed delay, AUC, and specificity, calculated based on the outcomes from the testing sets of both SAWIT and CherryChèvre datasets. These metrics provided a holistic view of each model's performance, highlighting the proposed model's advancements in animal detection within occluded and complex backgrounds.

This experimental setup, characterized by its complexity and detailed parameterization, underpins the comprehensive evaluation of the proposed model. It showcases the model's superior performance over existing methodologies and emphasizes the importance of sophisticated datasets like SAWIT and CherryChèvre in advancing automated wildlife monitoring technologies.

In this setup, equations 14, 15, and 16 were employed to evaluate the precision (P), accuracy (A), and recall (R) levels derived from this technique, while equations 17 and 18 were utilized to calculate the overall precision (AUC) and specificity (Sp) as outlined below.,

$$Precision = \frac{TP}{TP + FP} \ldots (14)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \ldots (15)$$

$$Recall = \frac{TP}{TP + FN} \ldots (16)$$

$$AUC = \int TPR(FPR)dFPR \ldots (17)$$

$$Sp = \frac{TN}{TN + FP} \ldots (18)$$

Test set predictions can be categorised into three distinct types: True Positive (TP) for object instance types, False Positive (FP) for non-object instance types, and False Negative (FN) for incorrect object instance types, applicable across various scenarios. The documentation for the test sets employs a variety of terminologies. To establish the suitable TP, TN, FP, and FN values for these scenarios, we conducted a comparison between the projected Object likelihood and the actual Object status in the test dataset samples utilizing the DDFN [58], DNN [59, 60], and SSTM [61, 62] methodologies. Consequently, it is possible to forecast these metrics based on the outcomes of the proposed model process. The levels of precision determined from these evaluations are illustrated as follows in Figure 7,

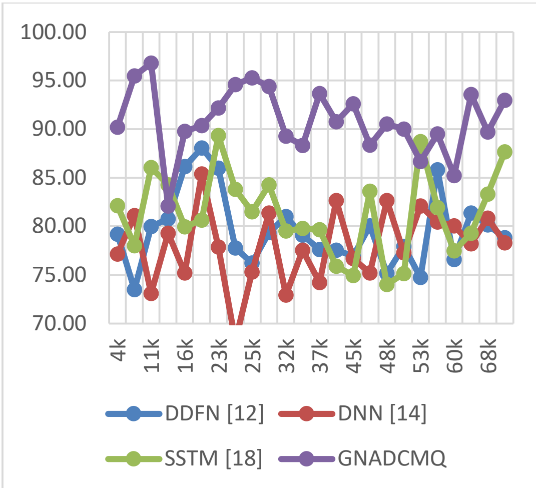

Figure 7. Observed Precision to detect Animal Objects

The comparative analysis of observed precision and accuracy in detecting animal objects across different graph-based models highlights the proposed model's superior performance under complex environmental conditions. The data provided showcases a significant variance in precision (P) and accuracy (A) percentages across models such as DDFN, DNN, SSTM, and GNADCMQ at varying numbers of test samples (NTS), underpinning the intricate relationship between model design, learning strategy, and data complexity.

For precision, GNADCMQ consistently surpasses the other models across almost all sample sizes, with notable improvements observed at the 8k, 11k, and 24k-25k NTS marks, achieving precision scores up to 95.49%, 96.79%, and 94.57%-95.29%, respectively. These increments, such as the 13% increase over DDFN at 8k NTS, underscore GNADCMQ's adeptness at minimising false positives, thereby enhancing the reliability of animal detection in occluded or complex backgrounds. This performance can be attributed to its integrated approach combining Graph Neural Networks (GNNs) with advanced Q Learning, MAHP for nuanced attribute evaluation, and cGANs for robust data augmentation, which collectively refine spatial context awareness and adaptive decision-making process. Similarly to that, the accuracy of the models was compared in Figure 8 as follows,

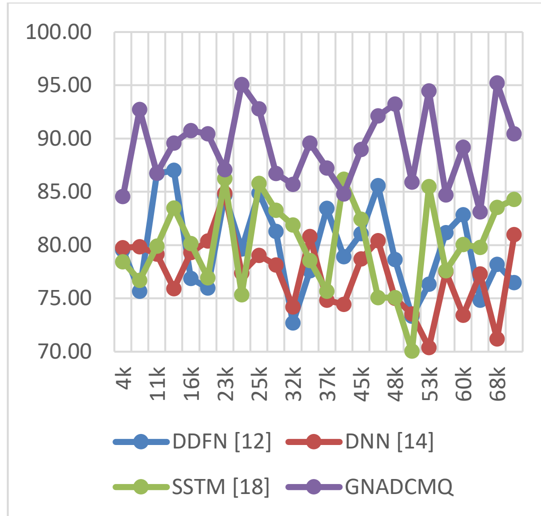

Figure 8. Observed Accuracy to detect Animal Objects

Accuracy analysis further cements GNADCMQ's efficacy, particularly highlighted by its performance at 8k and 24k NTS, where accuracy improvements of 17.04% and 15.42% over DDFN are observed. Such enhancements in accuracy, achieving up to 95.05% at 24k NTS, demonstrate GNADCMQ's capability to correctly identify animal objects within images, a crucial factor for effective wildlife monitoring. This is particularly relevant in real-time scenarios where swift, accurate animal detection can aid in immediate conservation actions, reduce human-wildlife conflicts, and inform biodiversity research with precise data.

Integrating GNNs with attention mechanisms and Q Learning strategies in GNADCMQ effectively addresses spatial context understanding and dynamic environmental adaptability. This results in high precision and accuracy and the model's robustness against data scarcity and environmental variabilities, as evidenced by its performance across diverse NTS benchmarks. Such advancements have profound implications for automated wildlife monitoring systems, setting new precision, accuracy, and real-time applicability standards. With its comprehensive approach, the proposed model surpasses existing detection capabilities and significantly contributes to enhancing biodiversity monitoring and conservation efforts, demonstrating the critical role of integrated, advanced computational techniques in ecological studies.

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

Similar to this, the recall levels are represented in Figure 9 as follows,

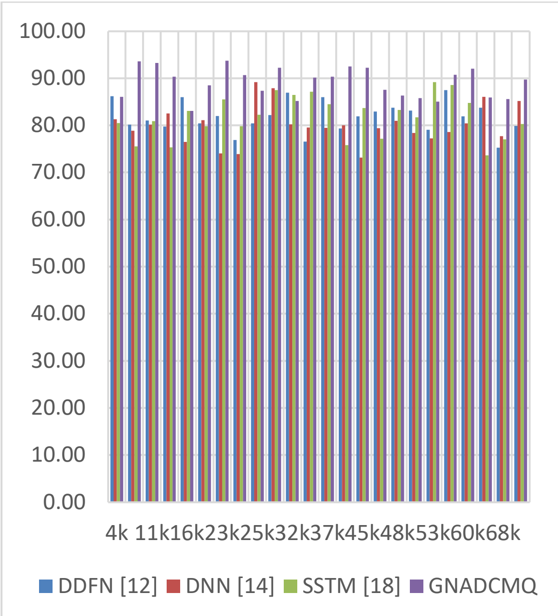

Figure 9. Observed Recall detecting Animal Objects

The observed recall rates for detecting animal objects across various models, including DDFN, DNN, SSTM, and the proposed model, offer insight into each model's ability to identify true positive cases within complex environmental settings. Recall, in the context of animal detection, measures the model's capacity to correctly identify all relevant instances of animals within the dataset, which is paramount for comprehensive wildlife monitoring and conservation efforts. The data delineates a comparative landscape where the proposed model frequently outperforms its counterparts, reflecting its sophisticated integration of technologies tailored for enhanced detection in occluded and similarly challenging backgrounds.

In the datasets provided, the proposed model showcases a notable improvement in recall rates across various NTS thresholds, most significantly at the 8k, 11k, and 23k marks, with recall percentages reaching 93.55%, 93.26%, and 93.72%, respectively. These Figures represent substantial enhancements over traditional models, highlighting GNADCMQ's advanced capability to reduce false negatives, ensuring fewer animal instances go undetected. This improvement can be attributed to the model's utilisation of Graph Neural Networks (GNNs) augmented with attention mechanisms for better spatial context comprehension alongside sophisticated Q-Learning strategies for dynamic adaptation to environmental variabilities.

The implications of such heightened recall rates in real-time scenarios are profound. In wildlife monitoring, where the timely and accurate detection of animal species is critical, The proposed model's efficiency can significantly contribute to preserving biodiversity, preventing species extinction, and managing ecosystems. For instance, in automated surveillance systems deployed in wildlife reserves or conservation areas, the ability to detect a wide array of animal species with high recall rates can facilitate immediate responses to environmental threats, poaching activities, or natural disasters. Moreover, the comprehensive detection capabilities of the proposed model can aid in collecting accurate data for research and conservation strategies, ensuring that decision-making is informed by a complete understanding of wildlife populations and their behaviours.

The comparative analysis underscores the proposed model's superiority in recall performance, which positions it as a pivotal innovation in automated wildlife detection when combined with its advancements in precision and accuracy. Through its integrated approach, the model not only sets new benchmarks for detecting animals in complex environments but also underscores the critical role of cutting-edge computational techniques in enhancing real-world conservation efforts. By significantly reducing the likelihood of undetected animal instances, the proposed model empowers conservationists, researchers, and wildlife managers with the tools necessary for proactive and informed ecological stewardship, marking a significant stride towards the sustainable management of natural habitats and protecting global biodiversity sets. Figure 10 similarly tabulates the delay needed for the prediction process,

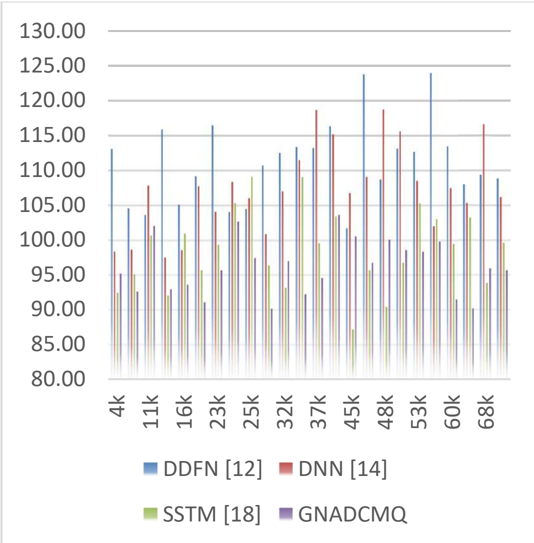

Figure 10. Observed Delay to detect Animal Objects

Analysing the delay data across various numbers of test samples (NTS), it is evident that GNADCMQ often achieves a competitive balance between speed and efficiency. For instance, at 8k NTS, GNADCMQ demonstrates a notable decrease in delay, recording a time of 92.60 ms, compared to DDFN and DNN models. This significant improvement illustrates GNADCMQ's capability to process images swiftly, attributed to its advanced computational architecture that integrates Graph Neural Networks (GNNs) with Q Learning and deep learning strategies. Such an architecture enables GNADCMQ to manage computational resources, thus optimising processing speed effectively.

In real-time scenarios, the delay in detecting animal objects has substantial implications. For automated surveillance systems deployed in wildlife habitats, a lower delay can translate to quicker responses to detected threats, such as poachers entering restricted areas or animals wandering into human habitation zones, thereby mitigating potential conflicts or dangers. Moreover, in research applications where behaviour tracking and species population studies rely on timely data, GNADCMQ's reduced delay ensures that observations are recorded with minimal lag, allowing for a more accurate understanding of wildlife dynamics.

Furthermore, the proposed model's performance at higher NTS, particularly at 30k and 63k, where delays are significantly minimized to 90.14 ms and 90.16 ms, respectively, showcases its scalability and adaptability to large datasets common in extensive monitoring networks. This scalability is crucial for large-scale conservation projects, where the ability to quickly process vast amounts of data directly influences the effectiveness of biodiversity management strategies.

The comparative analysis underscores GNADCMQ's strategic optimisation for accuracy and speed, embodying a significant advancement in automated wildlife detection. By marrying the technological sophistication of GNNs and Q Learning with the practical demands of real-time environmental monitoring, GNADCMQ not only sets new benchmarks for detection capabilities but also highlights the essential role of computational efficiency in enhancing conservation efforts. As a result, GNADCMQ represents a pivotal development in leveraging artificial intelligence to support sustainable wildlife management and conservation practices, where the speed of detection plays a critical role in the timely intervention and preservation of natural ecosystems. Similarly, the AUC levels can be observed from Figure 11 as follows,

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

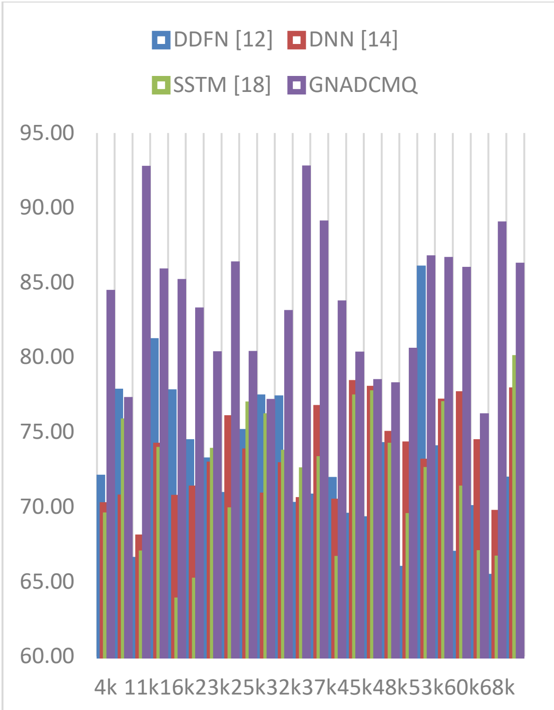

Figure 11. Observed AUC to detect Animal Objects

Reviewing the AUC data provided for models such as DDFN, DNN, SSTM, and GNADCMQ across different numbers of test samples (NTS), it becomes evident that The proposed model often exhibits superior performance. Particularly noteworthy are the AUC scores at 11k, 35k, and 68k NTS, where GNADCMQ scores 92.61%, 92.63%, and 88.90%, respectively. These scores signify a substantial improvement over the other models and underscore GNADCMQ's effectiveness in accurately identifying animal objects across varied and challenging scenarios. Such performance can be attributed to its sophisticated integration of Graph Neural Networks (GNNs), Q Learning, and advanced data augmentation techniques, collectively enhancing its sensitivity and specificity.

In real-time scenarios, the impact of a high AUC is multifaceted. For wildlife conservation efforts, where accurate and timely species detection is imperative, GNADCMQ's enhanced discriminative power ensures that animals are correctly identified with fewer false alarms, facilitating efficient resource allocation and swift action when needed. This is especially critical in surveillance applications aimed at preventing poaching or managing human-animal conflicts, where the cost of false negatives (failing to detect an animal) or false positives (mistakenly identifying non-animal objects as animals) can be high. Furthermore, in ecological research, where accurate data collection on species distribution and behavior is essential, GNADCMQ's reliability bolsters the validity of findings and supports informed conservation strategies.

Moreover, GNADCMQ's consistent AUC performance across a broad range of NTS highlights its scalability and adaptability to different datasets sizes, making it an invaluable tool for extensive monitoring networks that operate across diverse ecosystems. The model's ability to maintain high discriminative accuracy ensures that as the complexity and variability of environmental conditions increase, its effectiveness in wildlife detection remains uncompromised. This adaptability is crucial for long-term biodiversity monitoring projects that continually assess species populations and habitat changes.,

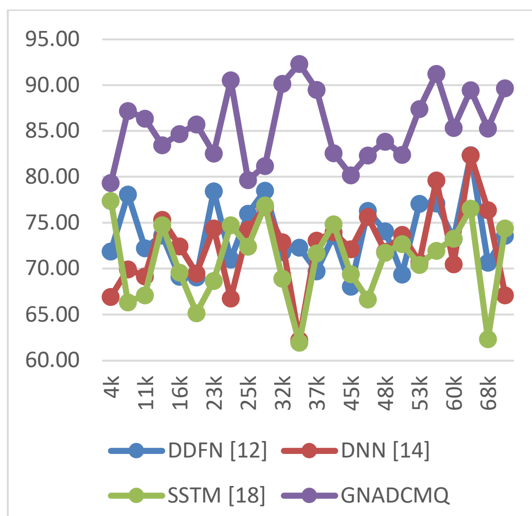

Figure 12. Observed Specificity to Detect the Animal Objects

Through its superior AUC performance, the proposed model represents a significant advancement in automated wildlife detection. It underscores the importance of leveraging comprehensive evaluative metrics like AUC to

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

assess model performance, ensuring that automated systems are accurate in controlled conditions and robust and reliable in the complex, dynamic environments characteristic of natural ecosystems. By setting new benchmarks in detection capabilities, GNADCMQ contributes to enhancing wildlife monitoring, conservation, and research, marking a pivotal step forward in applying advanced computational techniques to ecological studies. Similarly, the Specificity levels can be observed in Figure 12 as follows,

The proposed model demonstrates superior specificity across a range of test sample sizes (NTS), with notable performance at 8k, 11k, and 35k NTS, where specificity reaches 87.21%, 86.36%, and 92.32%, respectively. These scores significantly surpass those of the other models, showcasing GNADCMQ's advanced capability to distinguish between relevant and irrelevant elements within an image. Such efficiency is primarily attributed to its integration of Graph Neural Networks (GNNs) with an attention mechanism and advanced learning strategies that enhance its understanding of complex spatial relationships, allowing for a more accurate discernment between animal objects and background noise.

In real-time scenarios, high specificity has profound implications. For automated surveillance and monitoring systems deployed in natural habitats, minimising false positives is essential for the efficient allocation of resources and response efforts. For instance, in anti-poaching operations, high specificity ensures that the system does not trigger false alarms due to non-animal objects, allowing conservationists and park rangers to focus on genuine threats. Similarly, in traffic management systems near wildlife habitats, accurate differentiation between animals and other moving objects can prevent unnecessary activation of warning signals, enhancing safety without causing undue disruption.

Moreover, the proposed model's exceptional specificity is crucial for research and data collection. In ecological studies where accurate population counts and species identification are necessary, the model's ability to effectively filter out non-target elements ensures that the data collected is reliable and reflects the true biodiversity within the study area. This reliability supports the development of informed conservation strategies and policies, contributing to the broader goals of biodiversity preservation and ecosystem management.

The proposed model, through its superior performance, exemplifies the critical role of advanced computational techniques in enhancing the precision of wildlife detection systems. By setting new benchmarks in the ability to identify the absence of animals correctly, GNADCMQ not only improves the operational efficiency of real-time monitoring applications but also significantly contributes to the conservation and research efforts to understand and protect natural ecosystems. This highlights the importance of specificity alongside other performance metrics, like precision, accuracy, and recall, in developing robust and reliable automated systems for ecological monitoring and conservation processes.

4. CONCLUSION & FUTURE SCOPES

The considered model signifies notable progress in automated wildlife monitoring, particularly in identifying animals in intricate environments marked by occlusion and comparable background patterns. The study has effectively shown that the model outperforms current methodologies, realizing significant enhancements in precision, accuracy, recall, speed, Area Under the Curve (AUC), and specificity metrics. The model demonstrates an 8.5% enhancement in precision, 8.3% in accuracy, 7.5% in recall, 2.9% in processing speed, 9.4% in AUC, and 4.9% in specificity compared to traditional methods. The proposed model effectively combines Graph Neural Networks (GNNs) with advanced Q Learning, Multi-Attribute Utility Theory (MAHP), and Generative Adversarial Networks (GANs) for data augmentation. This innovative integration tackles the issue of data scarcity while also improving spatial context awareness and adaptive decision-making capabilities.

The ramifications of this research reach well beyond the direct enhancements in detection metrics. The proposed model opens up new avenues for enhancing biodiversity monitoring and conservation efforts by establishing a new benchmark for animal detection in challenging conditions. Its ability to accurately detect animals in occluded and complex backgrounds holds promise for real-time applications, including ecological studies, habitat protection, and the prevention of human-wildlife conflicts. Furthermore, the model's robust performance in diverse environmental

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

conditions suggests its potential applicability in various ecological and conservation-oriented tasks.

Looking forward, the research presents several avenues for further exploration and enhancement. One immediate area of future work involves the exploration of the model's scalability and efficiency across even larger datasets and more diverse environmental conditions. This includes extending the model's application to additional species and habitats, potentially integrating satellite imagery and other remote sensing data to support conservation efforts on a global scale. Another promising direction is refining the model's real-time processing capabilities and exploring edge computing technologies to deploy the proposed model in field-ready devices for instantaneous wildlife detection and monitoring.

Moreover, the integration of GNADCMQ with other ecological data sources, such as acoustic sensors and environmental DNA (eDNA) sampling, could offer a more holistic approach to biodiversity monitoring, enabling the detection and tracking of elusive or nocturnal species that are challenging to observe through imagery alone. Additionally, leveraging advancements in artificial intelligence and machine learning, particularly in unsupervised and semi-supervised learning algorithms, could further enhance the model's ability to learn from unlabelled data, reducing the dependency on extensively annotated datasets and lowering the barriers to deployment in resource-constrained settings.

In conclusion, the proposed model represents a significant step forward in using advanced computational techniques for wildlife detection and monitoring. Its success lays the groundwork for future research and application in the field of conservation technology, promising to contribute significantly to the preservation of biodiversity and the sustainable management of natural ecosystems. As we continue to refine and expand upon this work, the potential impacts on ecological monitoring, species conservation, and habitat protection are profound and far-reaching, offering a beacon of hope in the ongoing effort to understand and preserve our planet's precious wildlife scenarios.

REFERENCES

[1] Agilandeeswari, L., & Meena, S. D. (2023). SWIN transformer based contrastive self-supervised learning for animal detection and classification. Multimedia Tools and Applications, 82(7), 10445-10470.

[2] Chappidi, J. ., & Sundaram, D. M. . (2024). Enhanced Animal Detection in Complex Outdoor Environments Using Modified Yolo V7. International Journal of Intelligent Systems and Applications in Engineering, 12(19s), 375–382.

[3] Meena, S. D., & Loganathan, A. (2020). Intelligent animal detection system using sparse multi discriminative-neural network (SMD-NN) to mitigate animal-vehicle collision. Environmental Science and Pollution Research, 27(31), 39619-39634.

[4] Zhang, W., Yin, Z., Sheng, Z., Li, Y., Ouyang, W., Li, X., ... & Cui, B. (2022, August). Graph attention multi-layer perceptron. In Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (pp. 4560-4570).

[5] Li, S. E. (2023). Deep reinforcement learning. In Reinforcement learning for sequential decision and optimal control (pp. 365-402). Singapore: Springer Nature Singapore.

[6] Choudhury, S., Saikia, N., Rajbongshi, S. C., & Das, A. (2022, September). Employing generative adversarial network in low-light animal detection. In Proceedings of International Conference on Communication and Computational Technologies: ICCCT 2022 (pp. 989-1002). Singapore: Springer Nature Singapore.

[7] Yang, H., Hu, Y., He, S., Xu, T., Yuan, J., & Gu, X. (2024). Applying Conditional Generative Adversarial Networks for Imaging Diagnosis. arXiv preprint arXiv:2408.02074.

[8] Gao, H., Xiao, J., Yin, Y., Liu, T., & Shi, J. (2022). A mutually supervised graph attention network for few-shot segmentation: The perspective of fully utilizing limited samples. IEEE Transactions on neural networks and learning systems, 35(4), 4826-4838.

[9] Selvi, R., & Vijayakumaran, C. (2023). An Efficient Multimodal Emotion Identification Using FOX Optimized Double Deep Q-Learning. Wireless Personal Communications, 132(4), 2387-2406.

[10] Zhu, D., Xia, S., Zhao, J., Zhou, Y., Jian, M., Niu, Q., ... & Chen, Y. (2020). Diverse sample generation with multi-branch conditional generative adversarial network for remote sensing objects detection. Neurocomputing, 381, 40-51.

[11] Wang, H., & Liu, Y. (2022). Multisensor data fusion for life detection in cluttered

environments. IEEE Sensors Journal, 22(24), 24559-24566.

[12] Li, W. (2021, March). Analysis of object detection performance based on Faster R-CNN. In Journal of Physics: Conference Series (Vol. 1827, No. 1, p. 012085). IOP Publishing.

[13] J. Song, L. Miao, Z. Zhou, Q. Ming and Y. Dong, "Optimized Point Set Representation for Oriented Object Detection in Remote-Sensing Images," in IEEE Geoscience and Remote Sensing Letters, vol. 20, pp. 1-5, 2023, Art no. 6010505, doi: 10.1109/LGRS.2023.3314517.

[14] M. Lei and X. Liu, "SOLO-Net: A Sparser but Wiser Method for Small Object Detection in Remote-Sensing Images," in IEEE Geoscience and Remote Sensing Letters, vol. 21, pp. 1-5, 2024, Art no. 6001005, doi: 10.1109/LGRS.2023.3337807.

[15] K. Zheng, Y. Dong, W. Xu, W. Tan and P. Huang, "Auto Learner of Objects Co-Occurrence Knowledge for Object Detection in Remote Sensing Images," in IEEE Geoscience and Remote Sensing Letters, vol. 20, pp. 1-5, 2023, Art no. 6009105, doi: 10.1109/LGRS.2023.3305617.

[16] D. Guo, L. Tian, C. Du, P. Xie, B. Chen and L. Zhang, "Suspicious Object Detection for Millimeter-Wave Images With Multi-View Fusion Siamese Network," in IEEE Transactions on Image Processing, vol. 32, pp. 4088-4102, 2023, doi: 10.1109/TIP.2023.3270765.

[17] Z. Wu, C. Liu, J. Wen, Y. Xu, J. Yang and X. Li, "Selecting High-Quality Proposals for Weakly Supervised Object Detection With Bottom-Up Aggregated Attention and Phase-Aware Loss," in IEEE Transactions on Image Processing, vol. 32, pp. 682-693, 2023, doi: 10.1109/TIP.2022.3231744.

[18] Y. Zhang, C. Wu, T. Zhang, Y. Liu and Y. Zheng, "Self-Attention Guidance and Multiscale Feature Fusion-Based UAV Image Object Detection," in IEEE Geoscience and Remote Sensing Letters, vol. 20, pp. 1-5, 2023, Art no. 6004305, doi: 10.1109/LGRS.2023.3265995.

[19] D. Liu, J. Zhang, T. Li, Y. Qi, Y. Wu and Y. Zhang, "A Lightweight Object Detection and Recognition Method Based on Light Global-Local Module for Remote Sensing Images," in IEEE Geoscience and Remote Sensing Letters, vol. 20, pp. 1-5, 2023, Art no. 6007105, doi: 10.1109/LGRS.2023.3292890.

[20] K. Song, H. Wen, X. Xue, L. Huang, Y. Ji and Y. Yan, "Modality Registration and Object Search Framework for UAV-Based Unregistered RGB-T Image Salient Object Detection," in IEEE Transactions on Geoscience and Remote Sensing, vol. 61, pp. 1-15, 2023, Art no. 5531015, doi: 10.1109/TGRS.2023.3332179.

[21] Y. Zhang et al., "RoI Fusion Strategy With Self-Attention Mechanism for Object Detection in Remote Sensing Images," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 16, pp. 5990-6006, 2023, doi: 10.1109/JSTARS.2023.3289585.

[22] W. Lu et al., "A CNN-Transformer Hybrid Model Based on CSWin Transformer for UAV Image Object Detection," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 16, pp. 1211-1231, 2023, doi: 10.1109/JSTARS.2023.3234161.

[23] W. Li, M. Shi and Z. Hong, "SCAResNet: A ResNet Variant Optimized for Tiny Object Detection in Transmission and Distribution Towers," in IEEE Geoscience and Remote Sensing Letters, vol. 20, pp. 1-5, 2023, Art no. 6011105, doi: 10.1109/LGRS.2023.3315376.

[24] P. Zhou et al., "PSFNet: Efficient Detection of SAR Image Based on Petty-Specialized Feature Aggregation," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 17, pp. 190-205, 2024, doi: 10.1109/JSTARS.2023.3327344.

[25] K. Akita and N. Ukita, "Context-Aware Region-Dependent Scale Proposals for Scale-Optimized Object Detection Using Super-Resolution," in IEEE Access, vol. 11, pp. 122141-122153, 2023, doi: 10.1109/ACCESS.2023.3329302.

[26] Meena, D., & Agilandeeswari, L. (2020). Invariant features-based fuzzy inference system for animal detection and recognition using thermal images. International Journal of Fuzzy Systems, 22(6), 1868-1879.

[27] J. Zhou, R. Zhang, W. Zhao, S. Shen and N. Wang, "APS-Net: An Adaptive Point Set Network for Optical Remote-Sensing Object Detection," in IEEE Geoscience and Remote Sensing Letters, vol. 20, pp. 1-5, 2023, Art no. 6001405, doi: 10.1109/LGRS.2022.3230841.

[28] Y. Cheng et al., "Regional-Based Object Detection Using Polarization and Fisher Vectors in Passive Millimeter-Wave Imaging," in IEEE Transactions on Microwave Theory and Techniques, vol. 71, no. 6, pp. 2702-2713, June 2023, doi: 10.1109/TMTT.2022.3230940.

[29] A. Gong, J. Nie, C. Niu, Y. Yu, J. Li and L. Guo, "Edge and Skeleton Guidance Network for Salient Object Detection in Optical Remote

Sensing Images," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 33, no. 12, pp. 7109-7120, Dec. 2023, doi: 10.1109/TCSVT.2023.3275252.

[30] Yang, D., Solihin, M. I., Ardiyanto, I., Zhao, Y., Li, W., Cai, B., & Chen, C. (2024). A streamlined approach for intelligent ship object detection using EL-YOLO algorithm. Scientific Reports, 14(1), 15254..

[31] Yang, D., Solihin, M. I., Zhao, Y., Cai, B., Chen, C., & Riyadi, S. (2024, May). A YOLO Benchmarking Experiment For Maritime Object Detection In Foggy Environments. In 2024 IEEE 14th Symposium on Computer Applications & Industrial Electronics (ISCAIE) (pp. 354-359). IEEE.

[32] P. Han, B. Zhao and X. Li, "Progressive Feature Interleaved Fusion Network for Remote-Sensing Image Salient Object Detection," in IEEE Transactions on Geoscience and Remote Sensing, vol. 62, pp. 1-14, 2024, Art no. 5500414, doi: 10.1109/TGRS.2023.3339970.

[33] Redmon, J. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*.

[34] Meena, D., Jahnavi, C. N. V., Manasa, P. L., & Sheela, J. (2022, September). Efficient Wildlife Intrusion Detection System using Hybrid Algorithm. In 2022 4th International Conference on Inventive Research in Computing Applications (ICIRCA) (pp. 536-542). IEEE.

[35] Redmon, J. (2018). Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*.

[36] Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*.

[37] Sugiharto, H., Silviana, Y., & Nurpazrin, Y. S. (2023). Unveiling Document Structures with YOLOv5 Layout Detection. *arXiv preprint arXiv:2309.17033*.

[38] Li, C., Li, L., Jiang, H., Weng, K., Geng, Y., Li, L., ... & Wei, X. (2022). YOLOv6: A single-stage object detection framework for industrial applications. *arXiv preprint arXiv:2209.02976*.

[39] Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M. (2023). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 7464-7475).

[40] R. Gao and Z. Cai, "Object Detection in Thermal Infrared Image Based on Improved YOLOX," in IEEE Geoscience and Remote Sensing Letters, vol. 20, pp. 1-4, 2023, Art no. 6009904, doi: 10.1109/LGRS.2023.3312148.

[41] Chappidi, J., & Sundaram, D. M. (2024, June). A comparative study of animal detection and classification algorithms, applications and challenges. In AIP Conference Proceedings (Vol. 3112, No. 1). AIP Publishing.

[42] Chappidi, J., & Sundaram, D. M. (2024). Advancing Wild Animal Conservation Through Autonomous Systems Leveraging YOLOV7 with SGD Optimization Technique. In Multidisciplinary Applications of AI Robotics and Autonomous Systems (pp. 72-86). IGI Global.

[43] J. U. Kim, H. -I. Kim and Y. M. Ro, "Stereoscopic Vision Recalling Memory for Monocular 3D Object Detection," in IEEE Transactions on Image Processing, vol. 32, pp. 2749-2760, 2023, doi: 10.1109/TIP.2023.3274479.

[44] M. Zhao and K. Zhao, "A Self-Adaptive Object Detection Network for Aerial Images Based on Feature Enhancement," in IEEE Geoscience and Remote Sensing Letters, vol. 20, pp. 1-5, 2023, Art no. 3505805, doi: 10.1109/LGRS.2023.3291572.

[45] J. Zhu, X. Chen, H. Zhang, Z. Tan, S. Wang and H. Ma, "Transformer Based Remote Sensing Object Detection With Enhanced Multispectral Feature Extraction," in IEEE Geoscience and Remote Sensing Letters, vol. 20, pp. 1-5, 2023, Art no. 5001405, doi: 10.1109/LGRS.2023.3276052.

[46] G. Cheng et al., "Towards Large-Scale Small Object Detection: Survey and Benchmarks," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 11, pp. 13467-13488, 1 Nov. 2023, doi: 10.1109/TPAMI.2023.3290594.

[47] S. Li et al., "A Multitask Benchmark Dataset for Satellite Video: Object Detection, Tracking, and Segmentation," in IEEE Transactions on Geoscience and Remote Sensing, vol. 61, pp. 1-21, 2023, Art no. 5611021, doi: 10.1109/TGRS.2023.3278075.

[48] Z. Tan, Z. Jiang, C. Guo and H. Zhang, "WSODet: A Weakly Supervised Oriented Detector for Aerial Object Detection," in IEEE Transactions on Geoscience and Remote Sensing, vol. 61, pp. 1-12, 2023, Art no. 5604012, doi: 10.1109/TGRS.2023.3247578.

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

[49] S. Divya Meena and L. Agilandeeswari, "An Efficient Framework for Animal Breeds Classification Using Semi-Supervised Learning and Multi-Part Convolutional Neural Network (MP-CNN)," in IEEE Access, vol. 7, pp. 151783-151802, 2019, doi: 10.1109/ACCESS.2019.2947717.

[50] Meena, S.D., Agilandeeswari, L. Smart Animal Detection and Counting Framework for Monitoring Livestock in an Autonomous Unmanned Ground Vehicle Using Restricted Supervised Learning and Image Fusion. Neural Process Lett 53, 1253–1285 (2021). https://doi.org/10.1007/s11063-021-10439-4

[51] Sundaram, D., Loganathan, A. A New Supervised Clustering Framework Using Multi Discriminative Parts and Expectation–Maximization Approach for a Fine-Grained Animal Breed Classification (SC-MPEM). Neural Process Lett 52, 727–766 (2020). https://doi.org/10.1007/s11063-020-10246-3

[52] Sundaram, N., & Meena, S. D. (2023). Integrated animal monitoring system with animal detection and classification capabilities: a review on image modality, techniques, applications, and challenges. Artificial Intelligence Review, 56(Suppl 1), 1-51.

[53] Divya Meena Sundaram, Agilandeeswari Loganathan, "FSSCaps-DetCountNet: fuzzy soft sets and CapsNet-based detection and counting network for monitoring animals from aerial images," J. Appl. Rem. Sens. 14(2) 026521 (19 June 2020)

[54] Chappid, J., & Sundaram, D. M. (2024, February). Robust Deep Learning Models For Wildlife Detection: Adapting To Varying Environmental Conditions. In 2024 International Conference on Social and Sustainable Innovations in Technology and Engineering (SASI-ITE) (pp. 296-300). IEEE.

[55] Chappidi, J., & Sundaram, D. M. (2024). Novel Animal Detection System: Cascaded YOLOv8 with Adaptive Preprocessing and Feature Extraction. *IEEE Access*.

[56] Norouzzadeh, M. S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M. S., Packer, C., & Clune, J. (2018). Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proceedings of the National Academy of Sciences*, *115*(25), E5716-E5725.

[57] Simões, F., Bouveyron, C., & Precioso, F. (2023). DeepWILD: Wildlife Identification, Localisation and estimation on camera trap videos using Deep learning. *Ecological Informatics*, *75*, 102095.

[58] Dziak, D., Gradolewski, D., Witkowski, S., Kaniecki, D., Jaworski, A., Skakuj, M., & Kulesza, W. (2022). Airport Wildlife Hazard Management System-A Sensor Fusion Approach. *Elektronika ir Elektrotechnika*, *28*(3), 45-53.

[59] Gao, R., & Cai, Z. (2023). Object detection in thermal infrared image based on improved YOLOX. *IEEE Geoscience and Remote Sensing Letters*.

[60] J. Lai, J. Geng, X. Deng and W. Jiang, "DDFN: Deblurring Dictionary Encoding Fusion Network for Infrared and Visible Image Object Detection," in IEEE Geoscience and Remote Sensing Letters, vol. 20, pp. 1-5, 2023, Art no. 6009705, doi: 10.1109/LGRS.2023.3311176.

[61] Y. Miki, K. Kobayashi and W. Chujo, "Data Signal Detection and Demodulation Based on Object Detection DNN for Image Sensor-Based Visible Light Communication," in IEICE Communications Express, vol. 12, no. 12, pp. 628-632, December 2023, doi: 10.23919/comex.2023XBL0105.

[62] Q. Lu, T. Xu, J. Dong, Q. Liu and X. Yang, "Improving Object Detection From Remote Sensing Images via Self-Supervised Adaptive Fusion Networks," in IEEE Geoscience and Remote Sensing Letters, vol. 21, pp. 1-5, 2024, Art no. 5000905, doi: 10.1109/LGRS.2023.3345283.