1. Jayasheelan PALANISAMY, 2. S DEVARAJU

ORCID: 1. 0000-0002-4716-960X; 2. 0000-0003-3116-4772

OPEN ACCESS

DOI: 10.15199/48.2025.04.34

Quantum YOLOv8: advanced object detection for forest animal encroachment using customized EfficientNetB0 backbone with SPPF and FPN integration

Quantum YOLOv8: zaawansowana detekcja obiektów w celu wykrywania wtargnięć zwierząt leśnych przy użyciu dostosowanego szkieletu EfficientNetB0 z integracją SPPF i FPN

***Abstract**: The encroachment of wildlife into rural villages and roads during night time presents significant safety hazards to local communities. Night vision capabilities are crucial for addressing these challenges, as most wildlife intrusions occur under low-light conditions. To mitigate these risks, we introduce Quantum YOLOv8, an advanced variant of the YOLO object detection model. This model integrates a customized EfficientNet B0 backbone for efficient and accurate feature extraction, alongside a Feature Pyramid Network (FPN) to enhance multi-scale detection capabilities. Trained on carefully curated datasets of CCTV images capturing wildlife intrusions, Quantum YOLOv8 demonstrates exceptional performance in detecting animals under challenging lighting conditions. The model achieved a mean Average Precision (mAP) of 0.90 at 0.5 IoU, underscoring its effectiveness in mitigating human-wildlife conflicts in rural areas through reliable nighttime surveillance. The integration of robust night vision features is pivotal in ensuring the safety by providing accurate detection in environments where visibility is compromised.*

***Streszczenie:** Wkraczanie dzikich zwierząt na tereny wiejskie i drogi w nocy stwarza poważne zagrożenia dla bezpieczeństwa lokalnych społeczności. Możliwości widzenia w nocy są kluczowe dla rozwiązania tych wyzwań, ponieważ większość wtargnięć dzikich zwierząt ma miejsce w warunkach słabego oświetlenia. Aby złagodzić te ryzyka, wprowadzamy Quantum YOLOv8, zaawansowaną wersję modelu wykrywania obiektów YOLO. Ten model integruje dostosowany szkielet EfficientNet B0 w celu wydajnej i dokładnej ekstrakcji cech, wraz z siecią piramid cech (FPN) w celu zwiększenia możliwości wykrywania w wielu skalach. Quantum YOLOv8, szkolony na starannie wyselekcjonowanych zestawach danych obrazów CCTV rejestrujących wtargnięcia dzikich zwierząt, wykazuje wyjątkową wydajność w wykrywaniu zwierząt w trudnych warunkach oświetleniowych. Model osiągnął średnią precyzję średnią (mAP) 0,90 przy 0,5 IoU, podkreślając jego skuteczność w łagodzeniu konfliktów między ludźmi a dzikimi zwierzętami na obszarach wiejskich poprzez niezawodny nadzór nocny. Integracja solidnych funkcji widzenia w nocy ma kluczowe znaczenie dla zapewnienia bezpieczeństwa poprzez zapewnienie dokładnego wykrywania w środowiskach, w których widoczność jest ograniczona.*

Keywords: Quantum YOLOv8, EfficientNetB0, night vision object detection, animal tracking
Słowa kluczowe: Quantum YOLOv8, EfficientNetB0, detekcja obiektów noktowizyjnych, śledzenie zwierząt

Introduction

The increasing incidents of wildlife encroaching into rural areas have become a significant concern, as animals like tigers and elephants often enter villages and cross roads, posing serious safety risks to the local residents [1]. Addressing this challenge is crucial, and our research introduces an advanced computer vision solution aimed at mitigating these risks. We propose Quantum YOLOv8, a refined variant of the You Only Look Once (YOLO) model, specifically optimized for night vision applications [2]. This model utilizes an EfficientNet B0 backbone and a Feature Pyramid Network (FPN) to enhance real-time object detection, especially in low-light conditions. Our methodology is built on carefully curated datasets that capture instances of tigers and elephants intruding into human settlements [3]. These datasets, sourced from various open platforms like Kaggle, are chosen due to the frequent and often dangerous interactions these animals have with people in rural areas. The dataset, illustrated in Fig. 1, includes CCTV images that provide valuable insights into these human-wildlife interactions [4]. This selection emphasizes the importance of understanding and addressing the complexities of human-wildlife conflicts, particularly in the environments where night-time visibility is a factor. To tackle the challenge, we delve into the evolution of the YOLO model, introducing Quantum YOLOv8 [5]. This model incorporates the EfficientNet B0 backbone [6], known for its balance between performance and computational efficiency, and integrates an innovative Feature Pyramid Network (FPN) for multi-scale feature extraction [7]. The combination of these components results in a powerful and efficient framework capable of processing input data rapidly and accurately, even in low-light conditions. The following sections of the study provide an in-depth exploration of the YOLO algorithm's components, focusing on how the EfficientNet B0 backbone and FPN contribute to precise and quick object detection, with a particular emphasis on night vision [8]. Our experimental results highlight the effectiveness of Quantum YOLOv8 in object detection, particularly in scenarios that require high precision and recall, such as night-time surveillance [9]. The Precision-Recall Curve in the study illustrates the model's ability to maintain a strong balance between accuracy and false positive rates. This capability is crucial for practical applications where wildlife intrusions pose significant threats, especially during the night. Quantum YOLOv8 achieved an impressive 0.902 mAP@0.5 for all classes, underscoring its robustness and reliability in real-world scenarios. This research contributes to the advancement of computer vision technologies aimed at addressing critical challenges in human-wildlife interactions, with a special focus on enhancing safety in rural areas during low-light or night-time conditions.

The Following are the novelty of this article are:

1. The article introduces Quantum YOLOv8, designed to improve object detection in low-light conditions.
2. It uses the EfficientNet B0 backbone to achieve high accuracy with low computational cost.
3. The model helps detect animals like tigers and elephants in rural areas to enhance safety using night vision.

Related works

Suman Bhattachary et al. [10] discuss the urgent need to protect the Royal Bengal Tigers in the Sundarbans, the world's largest continuous mangrove forest. They propose using drones equipped with GPS and thermal detection systems to monitor and track tigers in remote and hard-to-reach areas while minimizing human interference. The drones will capture images of the tigers, which will be analyzed by a Deep Convolutional Neural Network (DCNN) utilizing YOLO and Faster R-CNN algorithms. This method aims to identify and count individual tigers by recognizing their unique features, thereby supporting conservation efforts. Shailendra Singh Kathait [11] discusses a method to classify images of tigers into individual classes using deep learning models. These classes represent the tigers themselves, with the aim of addressing the challenge of unique animal identification from the vast number of images captured by motion-activated cameras in natural habitats. Although existing techniques can identify animal species from images, distinguishing individual animals within a species remains difficult. The paper proposes a pipeline involving the YOLOv8 and EfficientNetB3 models with transfer learning to classify tiger images. From a dataset of 192 tigers, the study focused on 98 unique individuals, each represented by at least 15 images. To help prevent traffic accidents involving animals, Autoliv has developed an advanced night vision detection system for vehicles, already in use by Audi, BMW, and Daimler [12]. This system is the first of its kind available to consumers and uses a sophisticated classification method that can accurately detect animals even when they are partially hidden, in different poses, or at varying distances. It is supported by a large database containing thousands of hours of infrared video footage from around the world, with hundreds of thousands of images of animals in traffic situations. The system also features a tracking mechanism that monitors animal movements and predicts their behavior, along with a validation process that minimizes false alarms. Capable of detecting animals up to 200 meters away, the system uses special lights to highlight those that pose a potential danger, providing drivers with the information they need to respond quickly. This night vision system complements existing safety methods, giving drivers every opportunity to avoid accidents with animals. Many premium automotive brands now offer night vision systems designed to improve nighttime driving visibility. The latest generation of these systems often includes pedestrian detection features to help drivers avoid collisions. Yun Luo [13] examines pedestrian detection using two different night vision technologies: active night vision, which operates in the near-infrared (NIR) range, and passive night vision, which operates in the far-infrared (FIR) range. The study compares the advantages and disadvantages of each technology in terms of pedestrian detection effectiveness, collision avoidance, and market appeal. It also introduces an enhancement to the NIR active lighting approach that significantly improves pedestrian detection performance. With these advancements, the NIR night vision system is argued to be more effective at enhancing nighttime driving safety and may achieve broader market acceptance. Gabriel S. Ferrante et al [14] explores recent developments in animal detection and classification using computer vision technologies in urban areas, focusing on the studies published from January 2017 to May 2021. By searching two digital databases, 146 relevant studies were identified, and 20 were selected for a detailed review. These studies were grouped into six categories: (i) SVM, (ii) HOG, (iii) SIFT, (iv) PCA, (v) CNN, and (vi) DFDL. The analysis shows that CNN is the most frequently used method, and researchers are also experimenting with different combinations of techniques to boost accuracy. Despite these advancements, there remains a need for more specialized approaches to improve animal detection and classification in smart urban environments. The growing issue of animals entering human areas and illegal wildlife trafficking highlights the urgent need to protect our natural heritage. Monitoring wildlife at night is particularly challenging due to the high cost of night vision cameras. Nithya Madhasu [15] proposes a new approach to address these challenges. This approach combines night-time colorization using a Customized Conditional GAN (Generative Adversarial Network) with YOLO-CNAS (You Only Look Once-Customized Neural Architecture Search) for advanced wildlife detection. The colorization helps reveal important nocturnal behaviors, improving our understanding of wildlife. The deep learning algorithms used in this method can effectively detect animals and identify smugglers both day and night. The proposed solution has shown impressive accuracy improvements, increasing from 55.73% to 72.54% and finally reaching 94.67%. The model was trained and tested using a diverse dataset from sources like iNaturalist, Unsplash, and Pexels. This innovative approach offers promising solutions for better wildlife protection and conservation, helping ensure a safer coexistence between humans and animals.

Xinyi Jiang et al [16] explores the use of camera traps in wildlife surveys and biodiversity studies, emphasizing their ability to generate large volumes of images or videos depending on their activation method. While deep learning techniques have been suggested to automate wildlife identification in these images, reducing manual effort and speeding up analysis, there has been limited research validating and comparing different models in real-world field conditions. To address this, Jiang's study created a wildlife image dataset from the Northeast Tiger and Leopard National Park (NTLNP dataset) and evaluated the recognition performance of three widely-used object detection models: YOLOv5, Cascade R-CNN with HRNet32, and FCOS with ResNet50 and ResNet101. The study also compared the performance of these models when trained on day and night data separately versus combined. The results showed that models trained on both day and night data together performed well, with an average mean average precision (mAP) of 0.98 in animal image detection and 88% accuracy in animal video classification. Notably, the YOLOv5m model achieved the highest recognition accuracy, highlighting the potential of AI technology to help ecologists efficiently analyze large datasets and save time in wildlife monitoring.

Dataset collection and preprocessing techniques

Dataset collection

The Quantum YOLOv8 Night Vision Object Detection dataset was created using data from open-source platforms and YouTube videos, with a specific focus on tigers and elephants entering villages. The dataset is categorized into two primary groups: elephants and tigers. These animals were chosen due to their frequent presence in Indian villages, making them crucial subjects for studying and managing human-animal conflicts. The dataset provides valuable insights into the interactions between these large animals and human communities, supporting efforts to mitigate conflicts and enhance safety. Fig. 1 shows sample images from the dataset.

Fig. 1. Sample images from the collected dataset, sourced from open-source platforms. Classes A and B represent elephants and tigers, respectively.

Data Pre-Processing

In the data preprocessing phase, we standardized image sizes by resizing them to a uniform 640x640 resolution. We then applied various augmentation techniques to enhance dataset diversity. These techniques included shearing, cropping, rotating images by 90 degrees, scaling images by up to 20%, zooming in by up to 23%, and converting images to grayscale. Additionally, we applied random horizontal and vertical flips to further diversify the dataset. Following these augmentation steps, the dataset was divided as shown in Table 1: 70% for training, 20% for validation, and 10% for testing. This comprehensive approach significantly contributed to improving the model's accuracy in our analysis.

Table 1. Dataset Preprocessing into train the model

Class name	Training	Validation	Testing
Elephant	1300	500	275
Tiger	1300	500	203

Proposed Methodology

YOLO Model

YOLO (You Only Look Once) [17-19] combined with night vision technology offers a powerful solution for real-time object detection in low-light or nighttime conditions. Night vision systems enhance visibility in dark environments, enabling YOLO to effectively identify and track objects even in near-total darkness. By integrating infrared or thermal imaging with YOLO's advanced object detection algorithms, the system can detect and classify objects such as vehicles, animals, or intruders with high accuracy during nighttime. This combination is particularly valuable in applications like wildlife monitoring, security surveillance, and autonomous vehicles, where reliable object detection in low-light conditions is crucial. The synergy between YOLO's rapid detection capabilities and night vision technology provides a robust tool for enhancing safety and awareness in environments where visibility is limited.

Quantum You Only Live Once V8 (YOLOv8)

YOLOv8, the latest evolution of the You Only Look Once (YOLO) series, establishes a new benchmark in real-time object detection, with significant enhancements in speed, accuracy, and night-time performance. Its refined architecture enables rapid processing while maintaining high precision in identifying and classifying objects, even in low-light and nighttime conditions [20]. With advanced feature extraction, YOLOv8 excels at detecting small and complex objects, ensuring reliable performance in challenging environments. It operates efficiently across diverse hardware, from powerful GPUs to mobile devices. YOLOv8's capability to detect multiple objects simultaneously, with exceptional accuracy in both daylight and nighttime scenarios, makes it ideal for applications like autonomous driving, security surveillance, and real-time analytics. This enhanced performance solidifies YOLOv8 as a leading solution in computer vision, offering robust and dependable object detection for the modern era.

Fig. 2 Clearly represent Quantum YOLOv8, there are three main components: Backbone, Neck, and Head. The optimized Quantum YOLOv8 enhances feature extraction, layer optimization, and reduces computation, making it efficient for night vision applications [21]. The Backbone of Quantum YOLOv8 utilizes EfficientNetB0, which is specifically designed to balance model performance with computational efficiency. This is achieved through a unique combination of innovative techniques, including Mobile Inverted Bottleneck Convolution (MBConv) Blocks, significantly improving the model's efficiency.EfficientNet-B0 is designed to improve computational efficiency through four main steps: the Stem Layer, MBConv Blocks, Block Repetitions and Strides, and the Head Layer with an SPPF layer. The Stem Layer starts with a convolutional operation using 32 filters, followed by batch normalization and the Swish activation function, which work together to reduce the size of the image while increasing the number of channels [22]. The MBConv Blocks are crucial for capturing features at different scales. They begin by expanding the input to detect complex patterns, then apply depthwise convolutions [23]. After this, the feature maps are adjusted using a squeeze-and-excitation module, which fine-tunes them based on their importance. Finally, the output goes through global average pooling, and small fully connected layers to move residual connections.

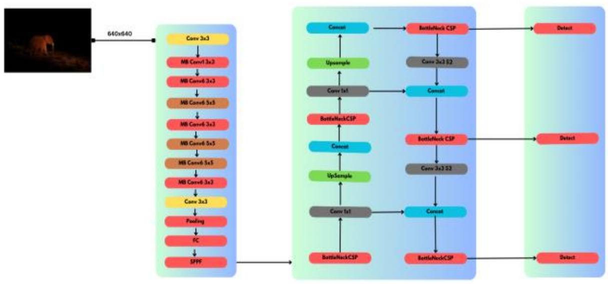

Fig. 2. Quantum YOLOv8 Architecture

$$X^{Final}=X+((z.ReLU6(X.W_{exp}.W_{dw})))$$

The consolidated MBConv block equation represents the sequence of operations within the block, starting with an input tensor *X*. The input is first expanded to a higher-dimensional space using a 1x1 convolution *Wexp*, followed by a depthwise convolution *Wdw* , which processes each channel independently. The X^{Final} output is then passed through the ReLU6 activation function and recalibrated using the Squeeze-and-Excitation (SE) module, where attention weights X are applied to the feature maps through element-wise multiplication$\odot$. The tensor is then projected back to the original dimensionality with another 1x1 convolution *Wproj*. Finally, if the input and output dimensions are the same, a residual connection adds the original input X to the processed tensor, forming the final output Xfinal. This equation encapsulates the process of expanding, filtering, and compressing the data while preserving important features through attention mechanisms and residual connections. The third main part of EfficientNetB0 is Block Repetitions and strides has a series of these MBConv Blocks with different configurations to make the process the input images with various expansion factors, Kernel size and strides. The Head layer of EfficientNetB0 is connected with the SPPF layer. In head efficient layer consists of conv2D, Global average pooling with fully connected layer Combination of Head layer with SPPF layer. In the neck of Proposed methodology is a Feature Pyramid Network (FPN) to enhance multi-scale feature extraction. The equation 3 FPN generates a series of feature maps that capture information across various resolutions by employing a top-down pathway. This pathway merges high-level semantic features with fine-grained details from higher-resolution layers. Consequently, the FPN facilitates comprehensive feature mapping, allowing for more effective detection of objects across different scales. The Quantum YOLOv8 model, which uses an EfficientNet B0 backbone, provides results with a list of bounding boxes that show where objects are detected in an image, along with their class labels and confidence scores.

This model is known for striking a great balance between speed and accuracy. It performs well in terms of mean average precision (mAP) at an intersection over union (IoU) of 0.5 and processes images quickly. Additionally, you can adjust the model's size to find the right balance between speed and accuracy without needing to retrain it, making it a flexible and effective choice for many uses.

Performance metric and Experimental results

Performance Metrics

F1 Confidence Curve

The F1 Confidence Curve depicts the relationship between F1 score and confidence thresholds in classification. It helps assess the precision-recall trade-off at various confidence levels.

$$F1ConfidenceCurve=2*(\text{ precision x recall}) /(\text{precision+recall})$$

Precision Confidence Curve

The Precision Confidence Curve shows how well a model performs at different confidence levels in classifying things. It helps us see how the precision, or accuracy, changes as we adjust confidence levels.

$$Precision=\text{True Positives}/(\text{True Positives+False Positives})$$

Recall Confidence Curve

The Recall Confidence Curve shows how the recall rate, which measures the ability to capture relevant instances, varies at different confidence thresholds in classification tasks. It visually depicts the changes in recall across various confidence levels, offering valuable insights into how well the model performs.

$$Recall=\text{True Positives}/(\text{True Positives+False Negatives})$$

Precision – Recall Curve

The Precision-Recall Curve shows how precision and recall are related at different decision thresholds in classification. It visually demonstrates how precision and recall values evolve with changing confidence levels.

Experimental Result

The experimental findings for P-YOLOv8 showcase its impressive ability to track animals. In Fig. 3 YOLOv8 model obtained a significant Mean Average Precision (mAP) of 0.93 at an intersection over union (IoU) threshold of 0.5 for object tracking. Fig. 4 illustrates the training and validation loss, along with the performance metrics of the model across different epochs. This includes the class loss, which plays a crucial role in classifying the target objects, among other losses. The figure demonstrates that the EfficientNet B0 backbone effectively reached an optimal local minimum during both the training and feature extraction phases.

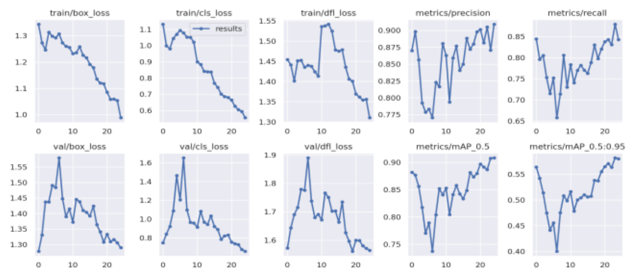

Fig. 3 Common Result for Quantum YOLOv8 algorithms

In Fig. 4, the F1 confidence curve and Recall confidence curve for the Quantum YOLOv8 model, trained specifically for a night vision dataset, are depicted. The wildlife animal classes include Elephant and Tiger. The F1 confidence curve shows that the F1 score for all class images is 0.87 at 0.274, representing the harmonic mean of precision and recall across all classes. Similarly, the Recall confidence curve indicates that the recall value for all classes in the Quantum YOLOv8 model is 0.98 at 0.000, clearly illustrating the impact of various threshold changes on precision and recall.

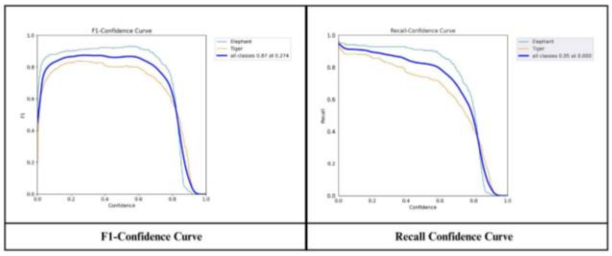

Fig. 4 F1 Confidence Curve and Precision Recall Curve for Quantum YOLOv8

Fig. 5 illustrates the Precision Confidence Curve and Precision-Recall Curve for the proposed surveillance and tracking system. In the Precision Confidence Curve, the model achieves an accuracy of 1.00 for all classes at a confidence threshold of 0.887, showcasing the model's precision. The Precision-Recall Curve depicts the relationship between precision and recall, represented through the Mean Average Precision (mAP). Notably, the figure emphasizes the Mean Average Precision for two crucial classes, with an overall mAP of 0.902 at an IoU threshold of 0.5, reflecting the model's high accuracy in surveillance and tracking functions.

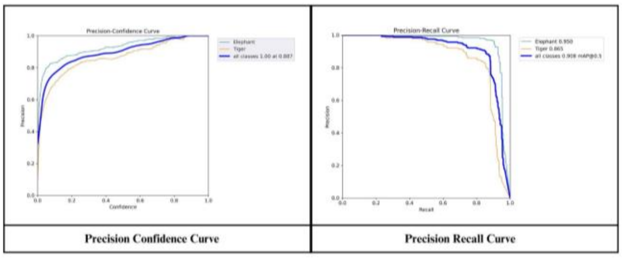

Fig. 5 Precision Confidence Curve and Precision Recall Curve for Quantum YOLOv8

Table 2 compares the performance of YOLOv5s, YOLOv7s, and Quantum YOLOv8 in detecting animals using night vision. The results show that Quantum YOLOv8 achieves higher accuracy and operates more efficiently than the other two models. This demonstrates that Quantum YOLOv8 is highly effective and holds great promise for improving animal detection in low-light environments.

Table 2 Performance comparison of the various YOLO algorithms

Model	Precision recall value	Recall confidence value	Accuracy
YOLOV5-s	0.83@mAP	0.82	82
YOLOV7-s	0.87@mAP	0.87	87
Quantum YOLOV8-s	0.902@mAP	0.90	90

Conclusion

The Quantum YOLOv8 model, designed with a focus on night vision, offers a substantial improvement in wildlife detection and safety in rural areas. By incorporating a customized EfficientNet B0 backbone and a Feature Pyramid Network (FPN), the model achieves higher accuracy and faster detection, especially in low-light scenarios where traditional models may underperform. With a mean Average Precision (mAP) of 0.902 at 0.5 IoU, Quantum YOLOv8 proves highly effective in identifying and tracking wildlife, such as tigers and elephants, in challenging environments. This enhanced detection capability is vital for reducing the risks associated with wildlife intrusions into human habitats, particularly during nighttime when these events are more common. The deployment of Quantum YOLOv8 can significantly contribute to the protection of rural communities, offering a dependable and efficient solution for managing human-wildlife interactions. This model's success underscores the potential for further innovation in computer vision technologies to improve safety and security in vulnerable areas.

Authors: *Jayasheelan Palanisamy, Bharathiyar University; S Devaraju, VIT University*

REFERENCES

[1]. Premarathna, K. S. P, Rathnayaka, R. M. K. T, Charles, J. An elephant detection system to prevent human-elephant conflict and tracking of elephant using deep learning. In *2020 5th IEEE International Conference on Information Technology Research (ICITR)* (2020, December), 1-6.
[2]. Sugumar, S. J, & Jayaparvathy, R. An improved real time image detection system for elephant intrusion along the forest border areas. *The Scientific World Journal*, 2014 (1), 393958.
[3]. Palanisamy, J, & Devaraju, S. Harmonizing Habitat: P-Yolov5 Enhanced Computer Vision For Mitigating Human-Wildlife Conflicts In Rural Areas. *Educational Administration: Theory and Practice*, (2024), 30(6), 2263-2272.
[4]. Amin, J, Shazadi, I, Sharif, M, Yasmin, M, Almujally, N. A., & Nam, Y. Localization and grading of NPDR lesions using ResNet-18-YOLOv8 model and informative features selection for DR classification based on transfer learning. *Heliyon*, (2024), 10(10).
[5]. Zhai, X, Huang, Z, Li, T, Liu, H, Wang, S. YOLO-Drone: an optimized YOLOv8 network for tiny UAV object detection. *Electronics*, (2023), 12(17), 3664.
[6]. Kansal, K, Chandra, T. B, Singh, A. ResNet-50 vs. EfficientNet-B0: Multi-Centric Classification of Various Lung Abnormalities Using Deep Learning Session id: ICMLDsE. 004. *Procedia Computer Science*, (2024), N.235, 70-80.
[7]. Chen, S, Zhao, J, Zhou, Y, Wang, H, Yao, R, Zhang, L, Xue, Y. Info-FPN: An Informative Feature Pyramid Network for object detection in remote sensing images. *Expert Systems with Applications*, (2023), N. 214, 119132.
[8]. Chen, S, Zhao, J, Zhou, Y, Wang, H, Yao, R, Zhang, L, & Xue, Y. Info-FPN: An Informative Feature Pyramid Network for object detection in remote sensing images. *Expert Systems with Applications*, (2023), N. 214, 119132.
[9]. Safaldin, M., Zaghden, N., & Mejdoub, M. An Improved YOLOv8 to Detect Moving Objects. *IEEE Access*, (2024).
[10]. Bhattacharya, S, Sultana, M, Das, B, Roy, B. A deep neural network framework for detection and identification of bengal tigers. *Innovations in Systems and Software Engineering*, (2024), 20(2), 151-159.
[11]. Kathait, S. S, Singh, V, Kumar, A. Individual Tiger Identification using Transfer Learning. *International Journal of Computer Applications*, N. 975, 8887.
[12]. Forslund, D, Bjärkefur, J. Night vision animal detection. In 2014 *IEEE intelligent vehicles symposium proceedings*, (2014), 737-742.
[13]. Y. Luo, J. Remillard and D. Hoetzer, Pedestrian detection in near-infrared night vision system, *IEEE Intelligent Vehicles Symposium, La Jolla,* CA, USA, 2010, 51-58, doi: 10.1109/IVS.2010.5548089.
[14]. G. S. Ferrante, F. M. Rodrigues, F. R. H. Andrade, R. Goularte and R. I. Meneguette, "Understanding the state of the Art in Animal detection and classification using computer vision technologies," 2021 *IEEE International Conference on Big Data (Big Data)*, Orlando, FL, USA, 2021, 3056-3065, doi: 10.1109/BigData52589.2021.9672049.
[15]. Madhasu, N., & Pande, S. D. Revolutionizing wildlife protection: a novel approach combining deep learning and night-time surveillance. Multimedia Tools and Applications, (2024), 1-35.
[16]. Tan, M, Chao, W, Cheng, J. K, Zhou, M, Ma, Y, Jiang, X, Feng, L. Animal detection and classification from camera trap images using different mainstream object detection architectures. *Animals*, (2022), 12(15), 1976.
[17]. Sapkota, R, Qureshi, R, Calero, M. F, Hussain, M, Badjugar, C, Nepal, U, Karkee, M. Yolov10 to its genesis: A decadal and comprehensive review of the you only look once series. *arXiv preprint arXiv:* (2024), 2406.19407.
[18]. Szymon CHERUBIN, Wojciech KACZMAREK, Michał SIWEK, YOLO object detection and classification using low-cost mobile robot, *Przegląd Elektrotechniczny*, 100 (2024), nr 9, 29-33
[19]. L.L. YIN, M.N.Shah ZAINUDIN, W.H.Mohd SAAD, N.A. SULAIMAN, M.I. IDRIS, M.R. KAMARUDIN, R. MOHAMED, M.S.J.A RAZAK, Analysis Recognition of Ghost Pepper and Cili-Padi using MaskRCNN and YOLO, *Przegląd Elektrotechniczny*, 99 (2023), nr 8, 92-97
[20]. Li, E, Wang, Q, Zhang, J, Zhang, W, Mo, H, Wu, Y. Fish detection under occlusion using modified you only look once V8 integrating real-time detection transformer features. *Applied Sciences*, (2023), 13(23), 12645.
[21]. Kamel Ibrahim, M. *Improving Object Detection using Enhanced EfficientNet Architecture* (Doctoral dissertation), (2023).
[22]. Praneeth, D, Kumar, N. S, Nagaraju, V. Enhanced Detection and Segmentation of Retinal Exudates in Diabetic Retinopathy using a Feature Pyramid Network with EfficientNet-B0 Encoder. *Indian Journal of Science and Technology*, (2024), 17(32), 3377-3387.
[23]. Batool, A, Byun, Y. C. Lightweight EfficientNetB3 model based on depthwise separable convolutions for enhancing classification of leukemia white blood cell images. *IEEE access*, (2023), N.11, 37203-37215.