EURASIAN PHYSICAL TECHNICAL JOURNAL

2025, Volume 22, No. 2 (52)

https://doi.org/10.31489/2025N2/121-132

Received: 25/03/2025 Revised: 19/05/2025 Accepted: 23/06/2025 Published online: 30/06/2025

Research Article

UDC: 537.632; 520.66; 519.876.5

THERMAL INFRARED OBJECT DETECTION WITH YOLO MODELS

Turmaganbet U.[1], Zhexebay D.[1], Turlykozhayeva D.[1]*, Skabylov A.[1], Akhtanov S.[2], Temesheva S.[1], Masalim P.[1], Tao M.[3]

[1]al-Farabi Kazakh National University, Almaty, Kazakhstan
[2]Unmanned Aerial Vehicle Laboratory, Scientific Research Institute of Experimental and Theoretical Physics, Almaty, Kazakhstan
[3]School of Electronics and Information, Northwestern Polytechnical University, Xian, China
*Corresponding author: dana.turlykozhayeva@kaznu.edu.kz

Abstract. *Object detection is a fundamental task in computer vision and remote sensing, aimed at recognizing and categorizing different types of objects within images. Unmanned aerial vehicle - based thermal infrared remote sensing provides crucial multi-scenario images and videos, serving as key data sources in public applications. However, object detection in these images remains challenging due to complex scene information, lower resolution compared to visible-spectrum videos, and a shortage of publicly available labeled datasets and trained models. This article introduces a Unmanned aerial vehicle - based thermal infrared object detection framework for analyzing images and videos in public applications and evaluates the performance of YOLOv8n/v8s, YOLOv11n/v11s, and YOLOv12n/v12s models in extracting features from ground-based thermal infrared images and videos captured by Forward-Looking Infrared cameras, as well as from unmanned aerial vehicle - recorded thermal infrared videos taken from various angles. The YOLOv8n/v8s, YOLOv11n/v11s, and the latest YOLOv12n/v12s models were deployed on a Raspberry Pi 5 using the OpenVINO framework. The successful deployment of these models, including the most recent version, demonstrates their feasibility for unmanned aerial vehicle-based thermal infrared object detection. The results show that YOLOv8 and YOLOv11 achieved high accuracy and recall rates of 93% and 92%, respectively, while the YOLOv12 model demonstrated good precision but comparatively lower performance in accuracy and recall, suggesting the possibility for further improvement.*

Keywords: object detection, YOLO models, Unmanned aerial vehicle, Forward-Looking Infrared cameras, thermal infrared images, Raspberry Pi 5.

1. Introduction

Object detection is a key task in computer vision, primarily focused on classifying and locating specific objects within an image. Remote sensing images, captured by various sensors on different platforms, often contain multiple objects at various scales, making them a rich source for object detection. Remote sensing has been applied to spaceborne, aerial, and ground-based platforms [1, 2]. Ground remote sensing systems, which use platforms such as high towers, vehicles, and ships, provide valuable optically labeled datasets. However, there are fewer thermal infrared (TIR) datasets available for detecting multiple objects at ground level. Unmanned aerial vehicle (UAV) platforms, on the other hand, can capture high-resolution thermal images and videos, compensating for the lack of high-resolution thermal data from satellites due to sensor limitations. With the rapid development of UAVs, the demand for efficient and effective detection algorithms is growing. UAV-based TIR images are now used in precision agriculture (PA), which helps

optimize farming practices by detecting variations in soil and crops. However, UAV TIR platforms generate large volumes of unlabeled data, creating challenges for the development of transfer techniques, algorithms, and detection applications based on ground-labeled data [1-9]. Unlike optical sensors, TIR sensors can capture images in both day and night conditions. As TIR technology continues to improve, it has found widespread use in applications like body temperature detection, traffic monitoring, and public health and safety, receiving significant attention. Despite this, object detection on UAV TIR images and videos remains relatively underdeveloped. Previous studies have focused on pedestrian detection using TIR images in ground settings [3], as well as the detection of ships [4], vehicles [5], thermal bridges in buildings [6], and electrical equipment [7]. For instance, in [8] authors compared the accuracy and performance time of various surveillance systems for detecting human presence using thermal imagery. In [9] authors developed a method for detecting vehicles with TIR images, which was successfully applied to traffic flow monitoring. They demonstrated that thermal image-based object detection could be effectively used in road traffic surveillance [9].

In general, object detection and UAV technologies are growing fields with diverse applications, not only in computer vision and deep learning but also in areas such as COVID-19 prevention and control, search and rescue operations, and Advanced Driver Assistance Systems (ADAS) using thermal imaging [10]. Despite these advances, object detection from UAV TIR images and videos continues to face numerous challenges, including complex backgrounds, low resolution, long imaging distances, and the presence of multiple scenes and objects [11-13].

Recently, YOLO-based object detection in thermal infrared imagery has gained significant traction, offering enhanced accuracy and efficiency in diverse real-world applications. In [11], the authors proposed enhancing search and rescue (SAR) missions using YOLOv8 for human detection in thermal imagery, achieving high precision with YOLOv8n at 90%. To expand object detection beyond the visible spectrum, in [12] authors proposed a TIR detection framework using YOLO models for Forward-Looking Infrared (FLIR) cameras. In [13], the authors propose an improved Mask-RCNN algorithm for UAV TIR video stream target detection, enhancing efficiency and reducing storage, but with potential limitations in complex environments and varying target sizes. In [14], the authors introduced ALSS-YOLO, a lightweight detector for TIR aerial images, achieving state-of-the-art wildlife detection performance on the BIRDSAI and ISODTIR UAV datasets. The above-mentioned articles [9-15] demonstrate relatively good efficiency in object detection, but there is potential to further improve metrics such as precision, accuracy, and recall.

This article presents an object detection framework based on thermal infrared (TIR) imaging using unmanned aerial vehicles (UAVs) for public applications. It evaluates the performance of the latest YOLOv12 models alongside YOLOv8n/v8s and YOLOv11n/11s. These models were tested on ground-based TIR images and videos captured by FLIR cameras under both bright and dark conditions, prior to being applied to UAV-captured TIR videos from various angles. The YOLOv8n/v8s, YOLOv11n/11s, and the latest YOLOv12n/v12s models were successfully deployed on a Raspberry Pi 5 using the OpenVINO (Open Visual Inference and Neural Network Optimization) framework, demonstrating their feasibility for UAV-based TIR object detection. We believe that the proposed integrated approach has strong potential to enhance real-time object detection performance in public applications.

2. Background

2.1 YOLO models

The YOLO architecture [16-18] consists of three core components: the backbone, the neck, and the head. The backbone is responsible for extracting features from the input image, while the neck and head process these features to generate predictions for bounding boxes, object classes, and confidence scores. The loss function integrates squared errors for bounding box coordinates, square-rooted differences for width and height, and classification errors, with weighted coefficients to emphasize both localization precision and object detection accuracy. Table 1 presents the structures of the YOLO models.

YOLOv8 focuses on a lightweight architecture with CSPNet, anchor-free design, and balanced performance between speed and accuracy. YOLOv11 improves upon YOLOv8 by introducing C3k2 blocks and spatial attention (C2PSA), enhancing computational efficiency and detection precision. YOLOv12 further refines the architecture with R-ELAN backbones, FlashAttention-driven area attention, and 7×7 separable convolutions, achieving higher accuracy, especially for small or occluded objects, while

maintaining real-time speeds. Each version builds on its predecessor to optimize object detection for diverse scenarios [19-21].

Table 1. Comparison of YOLOv8/11/12 structure

YOLO model	Backbone	Neck	Activation	Loss	Models
YOLOv8	CSPNet+CBS+C2f	FPN+PAN+SPPF	SiLU	Anchor-free approach	YOLOv8n/s/m/l/x
YOLOv11	CSPNet+C3k2	FPN+PAN+SPPF+C2PSA	LeakyReLU	Bounding Box,Class	YOLOv11n/s/m/l/x
YOLOv12	CSPNet+R-ELAN	FPN+PAN+Area Attention	LeakyReLU	Bounding Box,Class	YOLOv12n/s/m/l/x

2.2 Evaluation metrics

In this study, the performance of the YOLO models for object detection training and validation was assessed using precision, recall, accuracy, and FPS (Frames Per Second).

Precision measures the proportion of true positive samples out of all the predicted positive samples, while recall measures the proportion of true positive samples out of all the actual positive samples [22-24]. The calculations for precision and recall are provided in Eqs. (1) and (2):

$$precision = \frac{TP}{TP+FP}, \quad (1)$$

$$recall = \frac{TP}{TP+FN}, \quad (2)$$

TP (True Positive) refers to the number of persons that are correctly identified as persons, while FP (False Positive) indicates the number of instances where non-person objects are incorrectly classified as persons. FN (False Negative) represents the number of persons that are incorrectly classified as non-person objects.

Accuracy is a common performance metric in machine learning and object detection tasks, which measures how many predictions the model got right (both positive and negative) compared to all predictions made [25-28]. It is calculated using the formula:

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN}, \quad (3)$$

While accuracy provides a general sense of model performance, it can be misleading when there is a class imbalance. In such cases, other metrics like precision and recall are often used alongside accuracy to gain a more comprehensive understanding of the model's performance.

FPS is a performance metric used to measure how quickly a model can process and analyze video frames or images. In the context of object detection, FPS indicates how many frames the model can process in one second, giving an idea of the speed or real-time capability of the detection system. A higher FPS value means the model is able to process more frames per second, which is crucial for real-time applications such as video surveillance, autonomous vehicles, or drones. The ability to maintain high FPS while accurately detecting objects ensures that the system can operate effectively in dynamic environments where fast responses are necessary. In object detection tasks, balancing FPS with accuracy is important, as a model that runs too slowly might not be practical in real-time applications, even if it achieves high accuracy [29].

2.3 TIR image dataset

The FLIR Thermal Starter Dataset (FLIR, 2019) was used, which includes four object classes: person, car, bicycle, and dog. It contains both annotated thermal images and non-annotated optical images for training and validating object detection neural networks. The dataset comprises 10,288 annotated thermal infrared (TIR) images, featuring 28,151 persons and 46,692 cars, as well as 4,224 video-annotated TIR

images, with 21,965 persons and 14,013 cars. These images were captured on streets and highways in Santa Barbara, California, USA, between November and May, under mostly clear sky conditions, both during the day and at night. The images were taken with either a FLIR Tau2 camera (13 mm f/1.0, 45° horizontal field of view (HFOV) and 37° vertical field of view (VFOV)) or a Boson camera. While TIR images typically have lower spatial resolution, smaller signal-to-noise ratios (SNR), and fewer texture features compared to conventional optical images, they offer the advantage of being captured both day and night. In contrast, optical images can only be collected during the day unless artificial lighting is available. To take advantage of both all-day thermal infrared imaging and the features of optical images, it was assumed that the shape features learned by YOLO models from optical datasets [30] for detecting persons and cars would also apply to TIR images. This assumption provides a solid foundation for using pre-trained models to detect objects in thermal infrared images, offering a reasonable starting point for effective detection in this type of imagery [31-33].

2.4. Performance of YOLOv8/v11/v12

The Raspberry Pi 5 is a high-performance single-board computer designed for real-time processing. Equipped with a 64-bit Broadcom BCM2712 processor (Cortex-A76, up to 2.4 GHz) and a VideoCore VII GPU (800 MHz), it offers substantial improvements over its predecessors. With options for 4 GB or 8 GB LPDDR4X-4267 RAM, it efficiently manages multitasking and large datasets. The PCIe 2.0 interface facilitates external accelerator integration, boosting data processing capabilities. Despite requiring 5A and 5V of power, the Raspberry Pi 5 remains compact and efficient for various applications [34-37].

To deploy YOLOv8, YOLOv11 and YOLOv12 on the Raspberry Pi 5, the trained models were converted to an OpenVINO-compatible format. Using OpenVINO's Model Optimizer, the best.pt file was converted into an Intermediate Representation (IR) model, optimizing it for more efficient execution on resource-constrained devices. OpenVINO is an Intel toolkit that optimizes deep learning models for CPUs, GPUs, FPGAs, and VPUs. It converts models from frameworks such as TensorFlow and PyTorch into IR format using the Model Optimizer, and the Inference Engine ensures efficient execution. When integrated with the Raspberry Pi 5, OpenVINO improves inference efficiency, reducing computational overhead and enabling real-time processing through its BCM2712 processor and VideoCore VII GPU.

After converting the model to the OpenVINO format, the best_yolov8s_openvino_model was deployed on the Raspberry Pi 5. Test images stored in./test/images, along with their corresponding YOLO-format annotations in./test/labels, were used for evaluation. A Python script running on the Raspberry Pi 5 handles the complete object detection process: it loads and preprocesses images, performs inference using the YOLO model via OpenVINO, and compares the predicted bounding boxes with ground truth annotations. This workflow enables a detailed assessment of the model's accuracy and performance in real-time object detection on resource-constrained hardware.

For each test image, the following steps are performed: the image is loaded into memory, processed through the YOLO model to obtain predicted object coordinates, and compared with the ground truth annotations. Bounding boxes are visualized, with predictions shown in green and actual annotations in red, providing a clear indication of detection accuracy. The inference time for each image is recorded to evaluate real-time performance. Finally, the processed images, with bounding boxes, are saved in the./results directory for further analysis. This methodology ensures a thorough evaluation of both detection accuracy and processing efficiency.

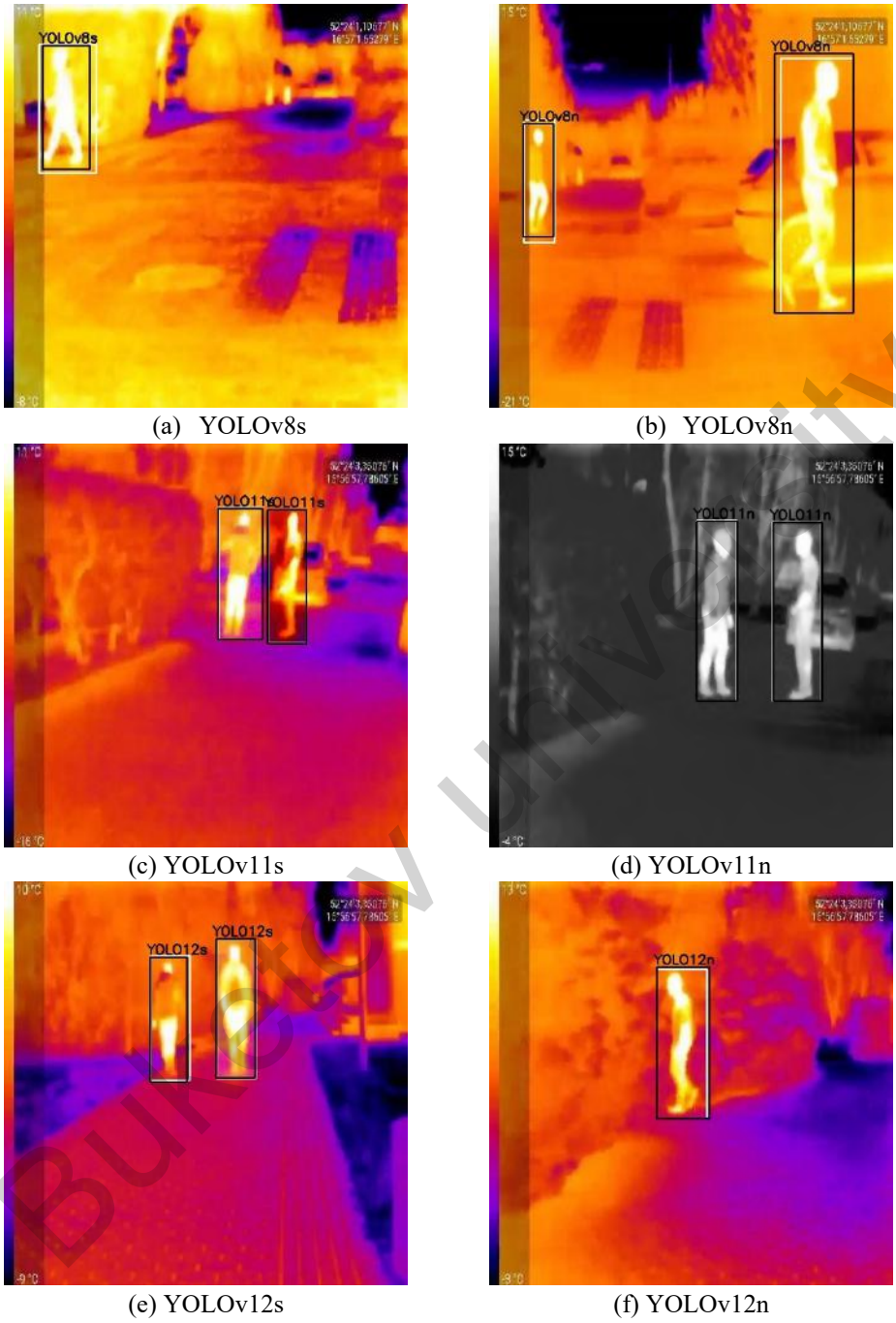
3. Results and discussions

3.1 Confusion matrix and accuracy

In Figure 1, an example of thermal reflections is shown, with thermal reflections placed in a box. To accurately identify the specific area of an object in a thermal image, it is crucial to detect and differentiate the thermal reflections. Figure 1 demonstrates that the detection capabilities of the YOLOv8, YOLOv11 and YOLOv12 models are suitable for use in a range of public applications and conditions. Figure 2 shows the confusion matrix (a, c) and accuracy curves (b, d) of the YOLOv8n/s models. The x-axis represents the true class labels of the samples, while the y-axis indicates the predicted results.

In Figure 2, YOLOv8n (a) correctly identifies 213 people with 23 background instances misclassified as people and 6 people as background, while YOLOv8s (c) detects 212 people correctly, misclassifying 23 background instances as people and 7 people as background. The accuracy curves in (b) and (d) illustrate the

training process over 100 epochs, showing a sharp initial improvement followed by a gradual stabilization around 92%.



(a) YOLOv8s (b) YOLOv8n

(c) YOLOv11s (d) YOLOv11n

(e) YOLOv12s (f) YOLOv12n

Fig.1. Thermal images with thermal reflections.

Figure 3 presents the results of the confusion matrix (a, c) and accuracy curves (b, d) of the YOLOv11n and YOLOv11s models. In Figure 3, YOLOv11n (a) accurately identifies 211 people, misclassifying 30 as background and 8 background instances as people, resulting in zero correctly classified background instances, while YOLOv11s (c) also detects 211 people correctly, misclassifying 21 as background and 8 background instances as people, again leading to no correctly identified background instances. The accuracy curves in (b) and (d) illustrate the training progression over 100 epochs, showing a steep initial gain before gradually stabilizing around 93%.

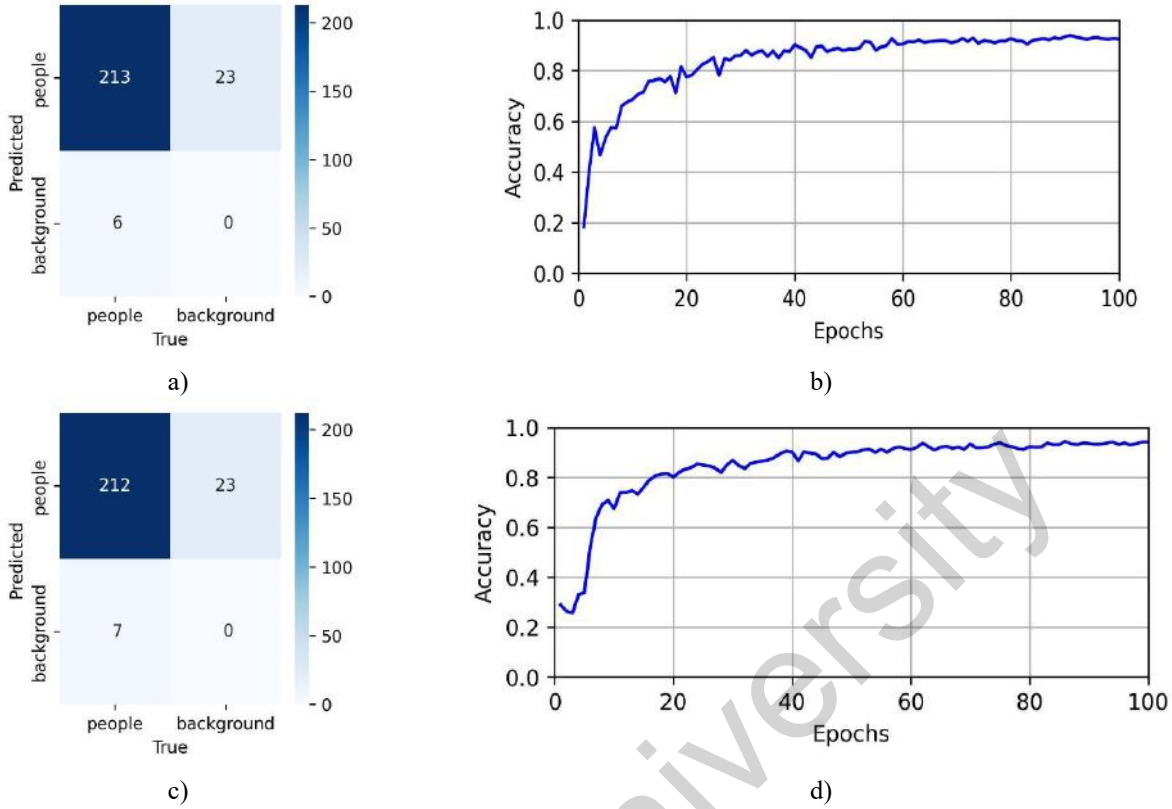


Fig. 2. (a, c) – YOLOv8n/s confusion matrix; (b, d) – Accuracy curve

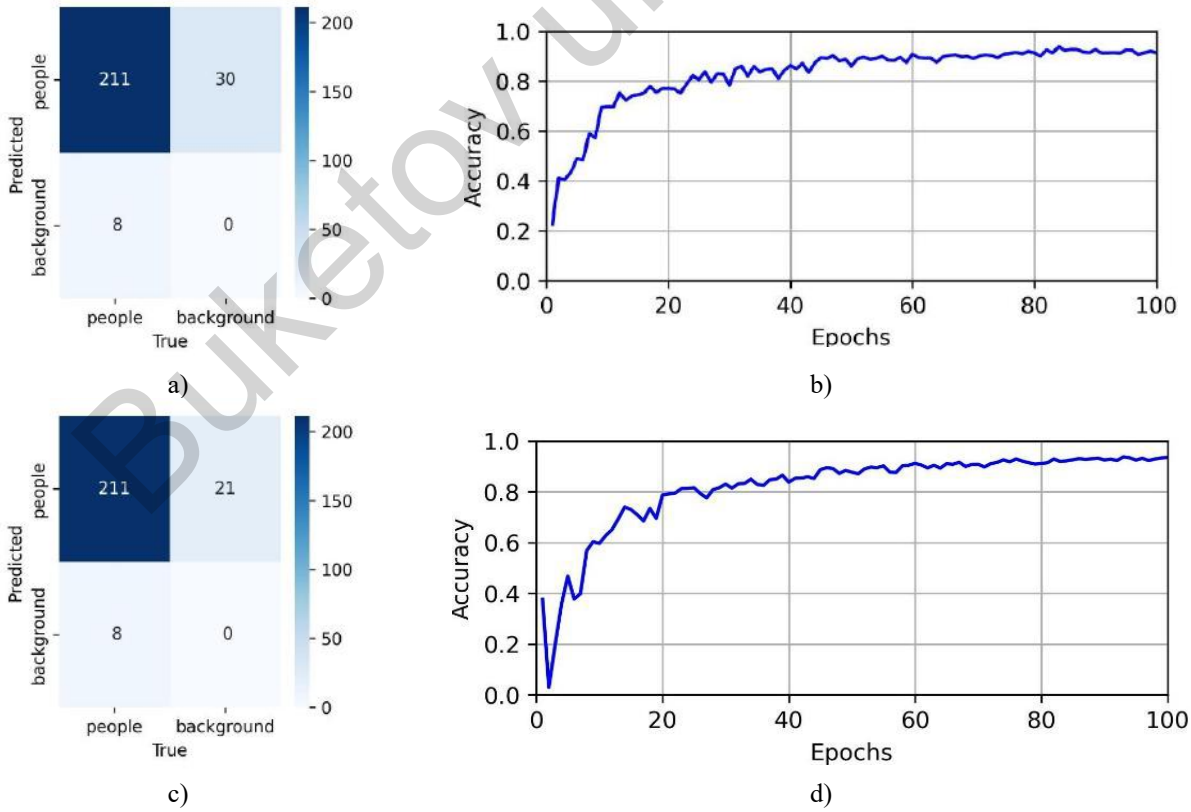


Fig.3. (a,c) – YOLOv11n/s confusion matrix; (b,d) – Accuracy curve

Figure 4 presents the results of the confusion matrix (a, c) and accuracy curves (b, d) of the YOLOv12n and YOLOv12s models.

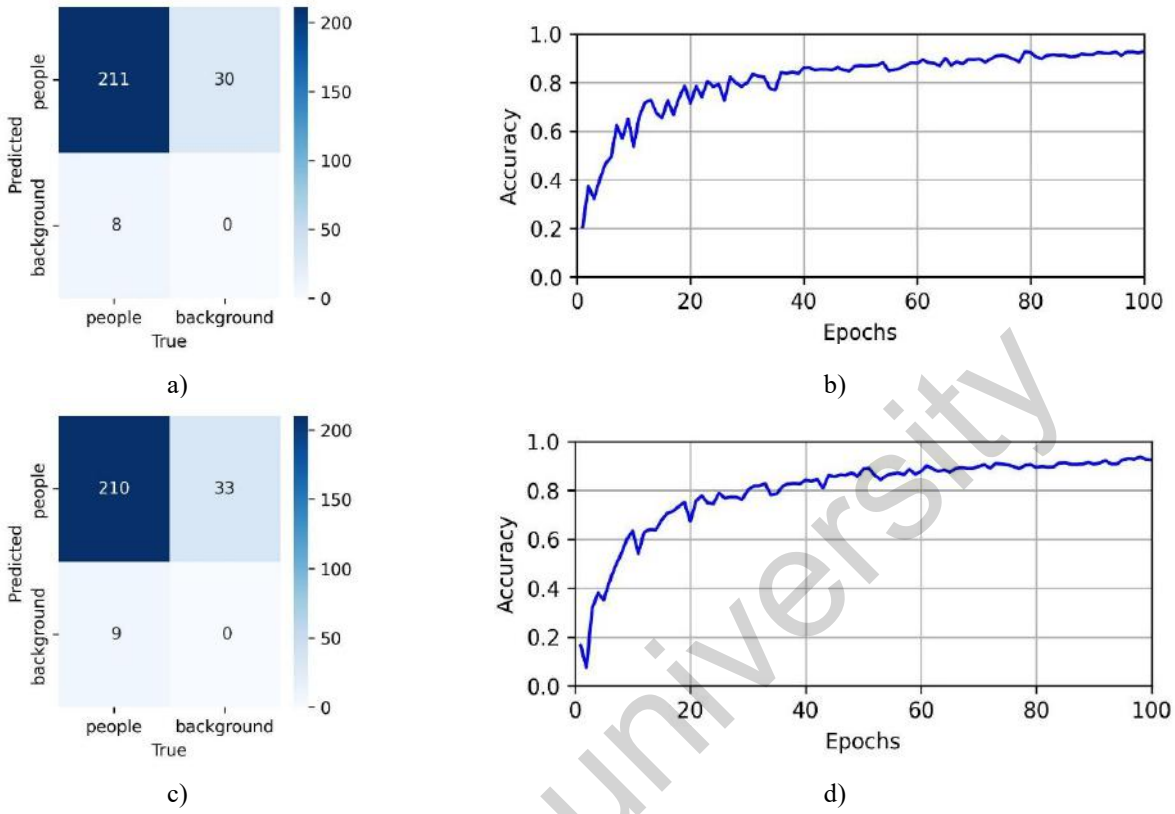


Fig.4. (a,c) – YOLOv12n/s confusion matrix; (b,d) – Accuracy curve

In Figure 4, YOLOv12n (a) correctly identifies 211 people, misclassifying 30 as background and 8 background instances as people, while YOLOv12s (c) detects 210 people, misclassifying 33 as background and 9 background instances as people, with both models failing to correctly classify any background instances. The accuracy curves in (b) and (d) depict the training process over 100 epochs, showing an abrupt initial gain before gradually stabilizing around 90%.

3.2 Recall and precision

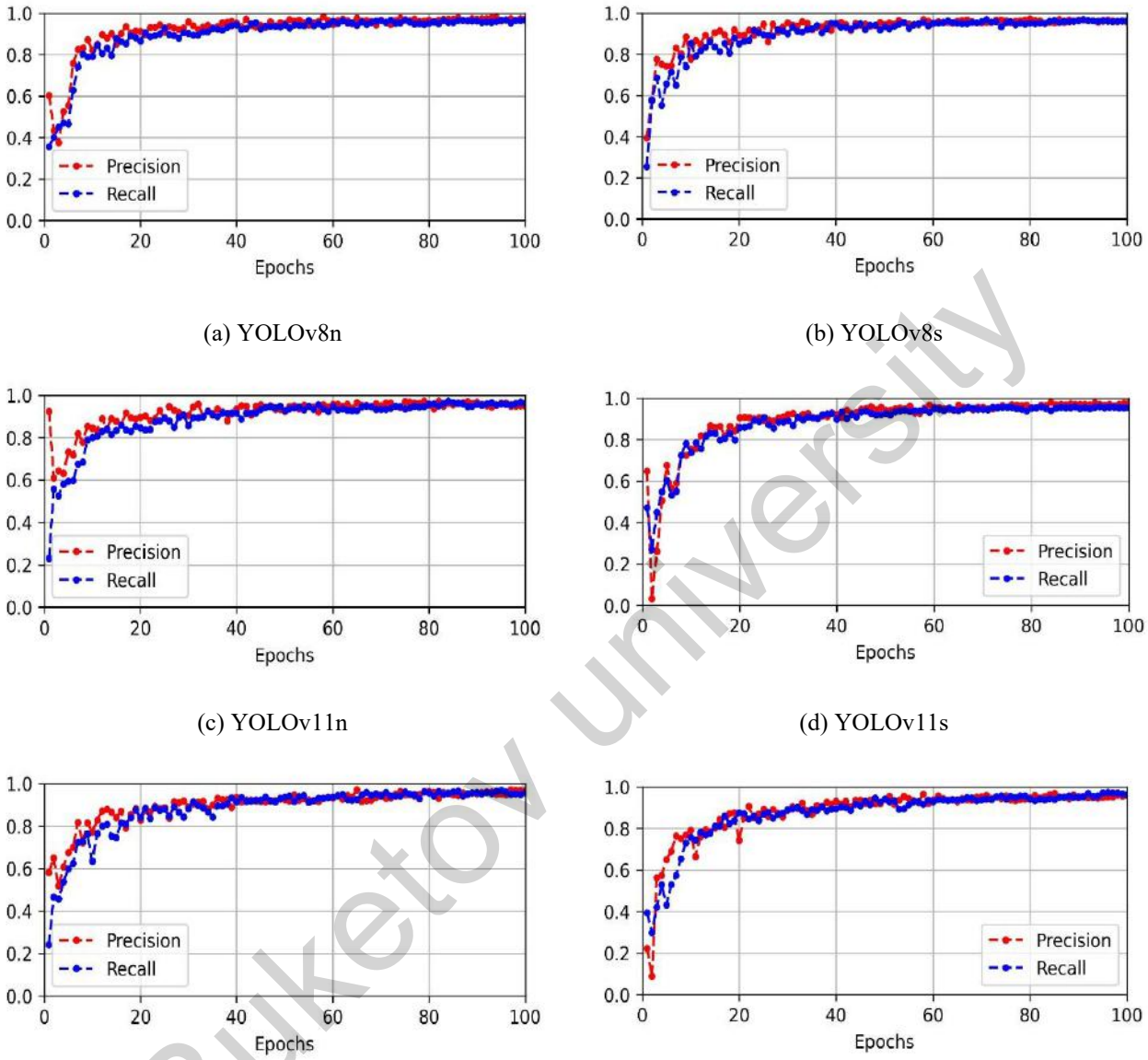
In Figure 5, the precision and recall values demonstrate the performance of YOLOv8n/s, YOLOv11n/s, and YOLOv12n/s over 100 training epochs. All models show rapid improvements in the first 20 epochs, followed by stabilization beyond epoch 50.

YOLOv8n/s begins with a Precision of approximately 0.60 and Recall around 0.55. By epoch 20, these rise to about 0.88 and 0.85, respectively, and eventually stabilize at around 0.95 (Precision) and 0.92 (Recall). YOLOv11n/s starts similarly with Precision near 0.58 and Recall at 0.50 but shows slightly more fluctuation during training. After minor instability around epoch 25, it reaches a stable Precision of about 0.94 and Recall of 0.91-0.92. YOLOv12n/s outperforms both, starting stronger with Precision around 0.62 and Recall at 0.58. It surpasses 0.90 in both metrics by epoch 20 and stabilizes after epoch 50 with Precision near 0.97 and Recall around 0.95. Its training curves are smoother and more stable, indicating better learning dynamics and generalization.

3.3 Accuracy and FPS comparison

Figure 6 (a, b) presents the comparative analysis of object detection accuracy and FPS for the YOLOv8n/s, YOLOv11n/s, and YOLOv12n/s models. According to Figure 6 (a), YOLOv8n and YOLOv8s achieve the highest accuracy, both nearing 0.88, indicating strong performance, while YOLOv11s also performs well at a similar level, suggesting optimizations in this version. YOLOv11n follows with slightly

lower accuracy, remaining above 0.84, while YOLOv12n is close behind. YOLOv12s records the lowest accuracy, approaching 0.82. According to Figure 6 (b), YOLOv11n achieves the highest FPS, exceeding 11, followed by YOLOv8n, which reaches approximately 10 FPS.



(a) YOLOv8n (b) YOLOv8s

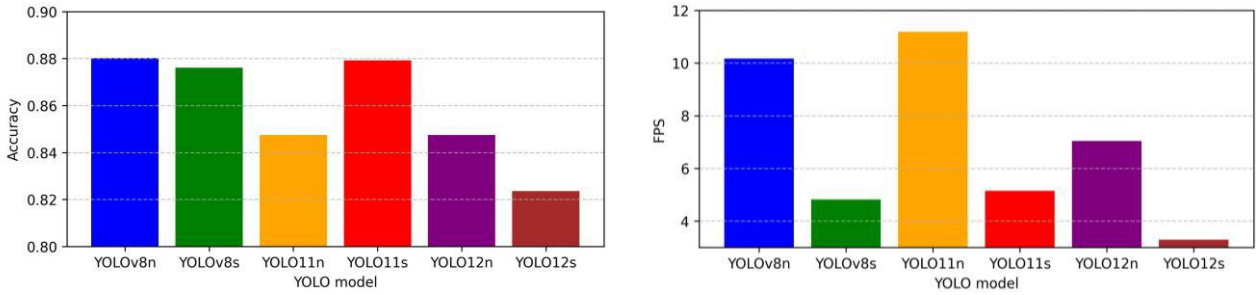
(c) YOLOv11n (d) YOLOv11s

(e) YOLOv12n (f) YOLOv12s

Fig.5. Precision and recall results of YOLO models.

YOLOv12n performs moderately, achieving approximately 7 FPS. YOLOv8s and YOLOv11s show lower performance, with FPS values around 5 and 4, respectively. YOLOv12s has the lowest FPS, measuring slightly above 2 FPS. Figure 7 provides a comparison of the efficiency of YOLOv8n/s, YOLOv11n/s and YOLOv12n/s models, showcasing their performance in terms of accuracy and processing speed.

Figure 7 demonstrates that YOLOv8n, YOLOv11n and YOLOv12n achieve the highest FPS and accuracy compared to the other evaluated models. Meanwhile, YOLOv8s and YOLOv11s exhibit slightly lower FPS but maintain competitive accuracy, balancing speed and precision in object detection, while YOLOv12s demonstrates the least efficiency.



(a) (b)

Fig. 6. Accuracy and FPS comparison of YOLO models

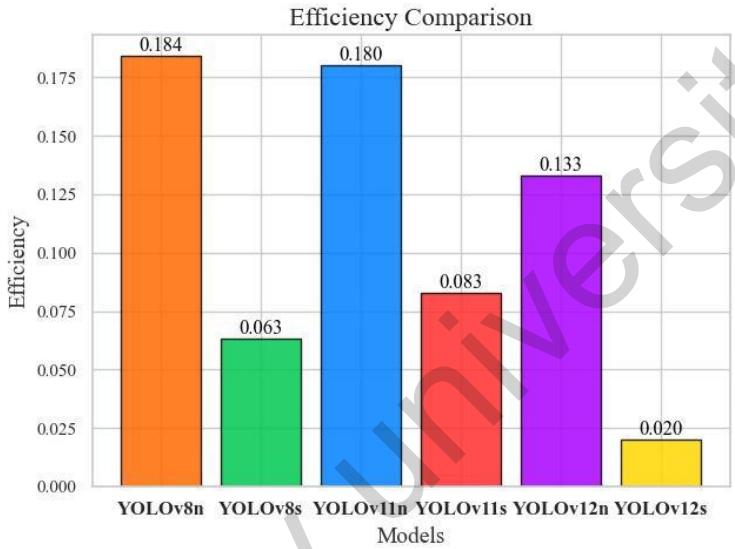


Fig.7. Efficiency Comparison of YOLOv8/11/12 models

4. Conclusion

In summary, this study assessed the performance of YOLOv8n/v8s, YOLOv11n/v11s, and YOLOv12n/v12s models for human detection in thermal infrared (TIR) UAV-captured imagery across various public applications. The results demonstrate that the YOLOv8n/v8s and YOLOv11n/v11s models offer notable improvements in detection accuracy (92% and 93%, respectively) and computational efficiency compared to YOLOv12n/v12s. YOLOv11n achieved the fastest detection speed at 11.50 frames per second, while YOLOv12s had the smallest model size. Additionally, the YOLOv11 models achieved the highest recall at 93% and strong precision of 92%. Their successful deployment on a Raspberry Pi 5 using the OpenVINO framework confirmed their feasibility for real-time object detection in thermal infrared imagery captured by UAVs.

The use of FLIR thermal infrared cameras provides a significant advantage by capturing high-quality imagery in both day and night conditions, allowing for effective object detection even in low-light and obscured scenarios. The ability to capture images from different angles further enhances detection performance, providing a broader view of the scene and improving object localization and identification in real-world applications. Future work will focus on integrating YOLO models with FPGA hardware and advanced thermal infrared imaging in UAV systems, enabling ultra-efficient, real-time object detection for public safety and monitoring applications. Additionally, we plan to utilize improved FLIR cameras with enhanced resolution and sensitivity to further boost detection accuracy and reliability during both day and night captures.

Conflict of interest statement

We want to make it clear that we have absolutely no conflicts of interest that could sway the finding or conclusions presented here. Financially, personally, or in terms of authorship, there's nothing that could interfere with the integrity of our work. It's important to us that our research is seen as unbiased and credible.

CRediT author statement

Turmaganbet U.K.: Conceptualization, Supervision, Zhexebay D.M.: Data Curation, Writing-Original Draft, Turlykozhayeva D.A.: Writing-Review&Editing, Skabylov A.A.: Data Curation, Writing-Original Draft, Akhtanov S.N.: Data Curation, Writing-Original Draft, Temesheva S.A.: Writing-Review&Editing, Masalim P.S.: Writing-Review&Editing, Tao M.: Writing-Review&Editing. The final manuscript was read and approved by all authors.

Funding

The research was funded by the Ministry of Science and Higher Education of Republic of Kazakhstan, grant AP19674715.

Acknowledgements

We would like to express our sincerest gratitude to the Research Institute of Experimental and Theoretical Physics of the al-Farabi Kazakh National University for supporting this work by providing computing resources of the Department of Physics and Technology for this study.

References

1 Ren X., Sun M., Zhang X., Liu L., Zhou H., Ren X. (2022) An improved mask-RCNN algorithm for UAV TIR video stream target detection. *International Journal of Applied Earth Observation and Geoinformation*, *106*, 102660. https://doi.org/10.1016/j.jag.2021.102660

2 Laghari A. A., Jumani A.K., Laghari R.A., Li H., Karim S., Khan A.A. (2024) Unmanned aerial vehicles advances in object detection and communication security review. *Cognitive Robotics*, https://doi.org/10.1016/j.cogr.2024.07.002

3 Mesvan N. (2021) Cnn-based human detection for uavs in search and rescue. *arXiv preprint arXiv:2110.01930*. https://doi.org/10.48550/arXiv.2110.01930

4 Jiang C., Ren H., Ye X., Zhu J., Zeng H., Nan Y., . Huo H. (2022) Object detection from UAV thermal infrared images and videos using YOLO models. *International Journal of Applied Earth Observation and Geoinformation*, *112*, 102912. https://doi.org/10.1016/j.jag.2022.102912

5 Haider A., Shaukat F., Mir J. (2021) Human detection in aerial thermal imaging using a fully convolutional regression network. *Infrared Physics & Technology*, *116*, 103796. https://doi.org/10.1016/j.infrared.2021.103796

6 Murthy J. S., Siddesh G.M., Lai W.C., Parameshachari B.D., Patil S.N., Hemalatha K.L. (2022) ObjectDetect: A Real-Time Object Detection Framework for Advanced Driver Assistant Systems Using YOLOv5. *Wireless Communications and Mobile Computing*, 2022(1), 9444360. https://doi.org/10.1155/2022/9444360

7 Yeom S. (2024) Thermal image tracking for search and rescue missions with a drone. *Drones,* 8(2), 53. https://doi.org/10.3390/drones8020053

8 Rizk M., Bayad I. (2025) Bringing Intelligence to SAR Missions: A Comprehensive Dataset and Evaluation of YOLO for Human Detection in TIR Images. *IEEE Access*. https://doi.org/ 10.1109/ACCESS.2025.3529484

9 Song Z., Yan Y., Cao Y., Jin S., Qi F., Li Z., Lu G. (2025) An infrared dataset for partially occluded person detection in complex environment for search and rescue. *Scientific Data,* 12(1), 300. https://doi.org/10.1038/s41597-025-04600-0

10 Dinh H.T., Kim E.T. (2025) A Lightweight Network Based on YOLOv8 for Improving Detection Performance and the Speed of Thermal Image Processing. *Electronics*, 14(4), 783. https://doi.org/10.3390/electronics14040783

11 Abbas Y., Al Mudawi N., Alabdullah B., Sadiq T., Algarni A., Rahman H., Jalal A. (2024) Unmanned aerial vehicles for human detection and recognition using neural-network model. *Frontiers in Neurorobotics*, 18, 1443678. https://doi.org/10.3389/fnbot.2024.1443678

12 Shao Z., Cheng G., Ma J., Wang Z., Wang J., Li D. (2021) Real-time and accurate UAV pedestrian detection for social distancing monitoring in COVID-19 pandemic. *IEEE Trans. Multimed.*, 1–16. https://doi.org/ 10.1109/TMM.2021.3075566

13 Shin D.-J., Kim J.-J. (2022) A deep learning framework performance evaluation to use YOLO in nvidia jetson platform. *Appl. Sci.*, 12 (8), 3734. https://doi.org/10.3390/app12083734

14 Jiang C., Ren H., Ye X., Zhu J., Zeng H., Nan Y., Huo H. (2022) Object detection from UAV thermal infrared images and videos using YOLO models. *International Journal of Applied Earth Observation and Geoinformation*, 112, 102912. https://doi.org/10.1016/j.jag.2022.102912

15 Zeng, S., Yang W., Jiao Y., Geng L., Chen X. (2024) SCA-YOLO: A new small object detection model for UAV images. *The Visual Computer*, 40(3), 1787-1803. https://doi.org/10.1007/s00371-023-02886-y

16 Kumar S., Yadav D., Gupta H., Verma O.P., Ansari I.A., Ahn C.W. (2021) A novel YOLOv3 algorithm-based deep learning approach for waste segregation: Towards smart waste management. *Electron,* 10, 1–20. https://doi.org/10.3390/electronics10010014

17 Li Z., Namiki A., Suzuki S., Wang Q., Zhang T., Wang W. (2022) Application of low-altitude UAV remote sensing image object detection based on improved YOLOv5. *Applied Sciences*, 12(16), 8314. https://doi.org/10.3390/app12168314

18 Liu P., Wang Q., Zhang H., Mi J., Liu Y. (2023) A lightweight object detection algorithm for remote sensing images based on attention mechanism and YOLOv5s. *Remote Sensing*, 15(9), 2429. https://doi.org/10.3390/rs15092429

19 Rasheed A.F., Zarkoosh M. (2025) Optimized YOLOv8 for multi-scale object detection. *Journal of Real-Time Image Processing,* 22(1), 6. https://doi.org/10.1007/s11554-024-01582-x

20 Rasheed A.F., Zarkoosh M. (2024) YOLOv11 Optimization for Efficient Resource Utilization. arXiv preprint arXiv:2412.14790. https://doi.org/10.48550/arXiv.2412.14790

21 Khanam R., Hussain M. (2024) Yolov11: An overview of the key architectural enhancements. arXiv preprint arXiv:2410.17725. https://doi.org/10.48550/arXiv.2410.17725

22 Jegham N., Koh C.Y., Abdelatti M., Hendawi A. (2024) Evaluating the evolution of yolo (you only look once) models: A comprehensive benchmark study of yolo11 and its predecessors. arXiv preprint arXiv:2411.00201. https://doi.org/10.48550/arXiv.2411.00201

23 Ussipov N., Akhtanov S., Turlykozhayeva D., Temesheva S., Akhmetali A., Zaidyn M., Tang X. (2024) MEGA: Maximum-Entropy Genetic Algorithm for Router Nodes Placement in Wireless Mesh Networks. *Sensors*, *24*(20), 6735. https://doi.org/10.3390/s24206735

24 Ussipov N., Akhtanov S., Zhanabaev Z., Turlykozhayeva D., Karibayev B., Namazbayev T., . Tang X. (2024) Automatic modulation classification for MIMO system based on the mutual information feature extraction. *IEEE Access*. 10.1109/ACCESS.2024.3400448

25 Andrushchenko M., Selivanova K., Avrunin O., Palii D., Tymchyk S., Turlykozhayeva D. (2024) Hand movement disorders tracking by smartphone based on computer vision methods. *Informatyka, Automatyka, Pomiary W Gospodarce I Ochronie Środowiska, 14: 5–10.* http://doi.org/10.35784/iapgos.6126

26 Turlykozhayeva, D., Temesheva, S., Ussipov, N., Bolysbay, A., Akhmetali, A., Akhtanov, S., & Tang, X. (2024, October). Experimental Performance Comparison of Proactive Routing Protocols in Wireless Mesh Network Using Raspberry Pi 4. In *Telecom* (Vol. 5, No. 4, pp. 1008-1020). MDPI. https://doi.org/10.3390/telecom5040051

27 Turlykozhayeva D.A., Akhtanov S. N., Baigaliyeva A.N., Temesheva S. A., Zhexebay D. M., Zaidyn M., Skabylov A.A. (2024) Evaluating Routing Algorithms Across Different Wireless Mesh Network Topologies Using Ns-3 Simulator. *Eurasian Physical Technical Journal*, *21*(2). https://doi.org/10.31489/2024No2/70-82

28 Turlykozhayeva D.A., Akhtanov S.N., Zhanabaev Z.Z., Ussipov N.M., Akhmetali A. (2025) A routing algorithm for wireless mesh network based on information entropy theory. *IET Communications*, *19*(1), e70011. https://doi.org/10.1049/cmu2.70011

29 Turlykozhayeva D., Waldemar W., Akhmetali A., Ussipov N., Temesheva S., Akhtanov S. (2024) Single Gateway Placement in Wireless Mesh Networks. *Physical Sciences and Technology*, *11*(1-2). https://doi.org/10.26577/phst2024v11i1a5

30 Turlykozhayeva D., Akhtanov S., Zhexebay D., Ussipov N., Baigaliyeva A., Wójcik W., Boranbayeva N. (2024) Evaluating machine learning-based routing algorithms on various wireless network topologies. Photonics Applications in Astronomy, Communications, Industry, and High Energy Physics Experiments, 13400, 236-245. https://doi.org/10.1117/12.3058676

31 Ibraimov M. K., Tynymbayev S. T., Skabylov A.A., Kozhagulov Y., Zhexebay D.M. (2022) Development and design of an FPGA-based encoder for NPN. *Cogent Engineering*, *9*(1), 2008847. https://doi.org/10.1080/23311916.2021.2008847

32 Zhexebay D., Skabylov A., Ibraimov M., Khokhlov S., Agishev A., Kudaibergenova G., Agishev A. (2025) Deep Learning for Early Earthquake Detection: Application of Convolutional Neural Networks for P-Wave Detection. *Applied Sciences*, *15*(7), 3864. https://doi.org/10.3390/app15073864

33 Ussipov N., Zhanabaev Z., Almat A., Zaidyn M., Turlykozhayeva D., Akniyazova A., Namazbayev T. (2024) Classification of Gravitational Waves from Black Hole-Neutron Star Mergers with Machine Learning. *Journal of Astronomy and Space Sciences,* 41(3), 149-158. https://doi.org/10.5140/JASS.2024.41.3.149

34 Akhmetali A., Namazbayev T., Subebekova G., Zaidyn M., Akniyazova A., Ashimov Y., Ussipov N. (2024) Classification of Variable Star Light Curves with Convolutional Neural Network. *Galaxies,* 12(6), 75.

AUTHORS' INFORMATION

Turmaganbet, Ulpan - PhD student, Department of Physics and Technology, al-Farabi Kazakh National University, Almaty, Kazakhstan; ORCID iD: 0009-0002-3782-425X ; uturmaganbet@gmail.com

Zhexebaym, Dauren – PhD, Researcher, Lecturer, Department of Physics and Technology, al-Farabi Kazakh National University, Almaty,Kazakhstan, ORCID iD: 0009-0008-1884-4662 ; zhexebay.dauren@kaznu.edu.kz

Turlykozhayeva, Dana - PhD, Researcher, Lecturer, Department of Physics and Technology, al-Farabi Kazakh National University, Almaty, Kazakhstan, ORCID iD: 0000-0002-7326-9196 ; turlykozhayeva.dana@kaznu.edu,kz

Skabylov, Alisher – PhD, Researcher, Lecturer, Department of Physics and Technology, al-Farabi Kazakh National University, Almaty, Kazakhstan, ORCID iD: 0000-0002-5196-8252 ; skabylov.alisher@kaznu.edu.kz

Akhtanov, Sayat - PhD, Researcher, Senior Lecturer, Unmanned Aerial Vehicle Laboratory, Scientific Research Institute of Experimental and Theoretical Physics, Almaty, Kazakhstan, ORCID iD:0000-0002-9705-8000; akhtanov.sayat@kaznu.edu.kz

Temesheva, Symbat - PhD student, Department of Physics and Technology, al-Farabi Kazakh National University, Almaty, Kazakhstan; ORCID iD: 0009-0000-2795-9586 ; symbat.temesheva@gmail.com

Masalim, Perizat – Master student, Department of Physics and Technology, al-Farabi Kazakh National University, Almaty, Kazakhstan, masalimperizat@gmail.com

Tao, Mingliang - PhD, Researcher, Professor, School of Electronics and Information, Northwestern Polytechnical University, Xian, China, ORCID iD: 0000-0002-0329-7124 ; mltao@nwpu.edu.cn