Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 16th International Scientific and Practical Conference. 2025. Volume V, 201-207

Examining the Impact of Image Augmentation on Object Detection and Classification to Improve Performance of Situational Awareness Systems

Svilen Neykov
Bulgarian Defence Institute ”Prof. Tsvetan Lazarov”
Sofia, Bulgaria
ORCID: 0009-0005-3975-2781
WoS: KEH-5558-2024
svilen@nkv.me

***Abstract*—Situational Awareness (SA) and Artificial Intelligence (AI) go hand-in-hand in future applications, be that civilian applications, disaster relief or especially in the military domain. There, providing the most information, as quickly and as precisely as possible is crucial to completion of the given mission both in acquiring and denying the materiel and personnel. Object detection and classification are integral parts of this and will be the cornerstone in the coming years for developments and advances in civilian and military life. it is vital that these systems can operate in all weather conditions and be able to discern subtle differences in the observed objects, without relying on additional equipment.**

Keywords— Artificial Intelligence, Object Detection, Situational Awareness.

I. INTRODUCTION

There are various fields where AI algorithms may enhance both situational awareness and decision-making speed and precision. Such may include disaster mitigation and remediation, search and rescue operations, vehicle and machinery operation, etc. Even more in the military where AI has not yet entered with its full force and hype. One key area here can be Target Acquisition and Friend-Foe identification, especially in theatres where similar materiel is used and there is a lack of training material for the algorithms. Here state-of-the-art Object detection algorithms come into play. In this paper one such algorithm is examined, both in a pre-trained and trained on a very limited dataset, as well as using augmentation techniques to see the impact it can have on results.

II. MATERIALS AND METHODS

A. Object Detection Algorithms

The human is a visual animal, most of the information our brains process is visual, and we have built the world around us based on this. Visual cues, movement, patterns and imperfections are a small part of the information we can process naturally. However, we are prone to boredom, tiredness, loss of vision due to various medical, psychological and environmental reasons. Here enters the digital eye. Sensors, optics that can be designed to be vastly superior to the human eye. They can process information faster, can operate in various environments and weather conditions, and work in light-bands impossible for the human eye. Thus, computer vision is introduced.

Computer Vision has wide application in almost every aspect of the modern world. Industrial applications such as defect monitoring, assembly verification, safety checks, maintenance, etc. Research applications such as identification of cell types, quality checks of drug crystalline structures, etc. In everyday life such as vehicle monitoring for toll collection, CCTV monitoring for law-enforcement, disaster prevention and relief, access verification for home and office spaces and so on.

One of the greatest fields where Computer Vision, can be used is n the military domain. Here information is vital. Having as much information as possible and being able to process this information as fast as possible and as accurately as possible provides an edge for you and your allies and puts the opposing forces at a measurable disadvantage. One of the major computer vision aspects is Object Detection and Classification. Traditionally object detection has utilized sliding windows which select regions of the image, but it is very hard to find the optimal number

Online ISSN 2256-070X
https://doi.org/10.17770/etr2025vol5.8491

of these windows. The traditional methods rely heavily on manually created features, which struggle with accuracy and robustness when viewing diverse objects and possess complex backgrounds. These features also have one additional flaw - they are not universal, this translates to poor multi-domain performance and high complexity and time-consumption in region selection. With Deep Learning (DL) booming in the industry, object detection utilizing DL has vastly improved and surpassed traditional object detection methods. DL-based algorithms use very intricate and strictly hierarchical architectures, these enable the training of more complex features. Also, owing to the significant learning capabilities of convolutional neural networks (CNNs) feature can be better represented than traditional methods. This in turn raises significantly the robustness and eliminates the manually created features. [1]

B. Situational Awareness

Situational Awareness refers to being aware of one's surroundings and reacting appropriately to identified stimuli. SA is important in all aspects of life, industrial production, emergency and disaster response, infrastructure monitoring, etc. It is however vitally important in the Military domain. SA is viewed as an integral part of the Modern Command and Control (C2) system alongside planning, tasking and control. Recently, a lot of research has been done on crisis assessment [2] [3], as well as C2 system models [4]. For a field commander or a chief of an emergency response center being aware of all possible actors on the field, their identification, intention, status and movement vector is crucial in developing a complete picture. This picture directly impacts the plan-making and decision process used. The SA and decision-making model which here we can use is based on the one proposed by Munir, Arslan, Alexander Aved, and Erik Blasch. Here we can see clearly 3 Levels of SA: Level 1 - Perception: here are listed items pertaining to identifying the presence of an object. Level 2 - Comprehension: here we can see the linking of the objects from level 1 into a more complete picture, understanding their relation and significance through the lens of the operators' objectives. Level 3 - Projection: here we can find the projection of possible future actions of the identified objects and possible ways to perform actions in a way that is outside of their SA. When discussing this model, we can see that the work presented is combining Level 1 and Level 2, as apart from identifying if there is an object, it tries to identify if this object is in the blue or red team, and so it enhances the understanding of the objects after identifying them. [5]

C. Target Acquisition and Friend-Foe Identification

Target Acquisition in the post-Cold War era refers to Identification and Assessment of capabilities, assets and identities of formations, systems and materiel in general. This is especially important in Identifying and Acquiring High Value Targets and High Pay-Off Targets, which when neutralized would provide a significant contribution to the success of the friendly/blue team forces. This process requires a significant structure model to have a robust process and possess the ability to choose the high-value targets (HVTs) and/or high-priority targets (HPTs). To facilitate this a friend-for identification system can also be used. This would lessen the computational burden by sifting through all the initial contacts and marking blue (friendly) forces beforehand. This type of system however has its limitations, as only friendly/blue forces may be positively identified, however they may not be able to respond, due to various reasons, some of which are malfunctions, lack of Identification friend or foe (IFF) equipment and so on. IFF is a tool within the broader military action of combat identification (CID), the characterization of objects detected in the field of combat sufficiently accurately to support operational decisions. The broadest characterization is that of friend, enemy, neutral, or unknown. CID not only can reduce friendly fire incidents but also contributes to overall tactical decision-making. [6]

Here is where additional Identification, mainly visual can have a massive positive effect and Artificial Intelligence, especially object detection and classification algorithms has its place.

III. Results and Discussion

A. Setup

Utilizing a state-of-the-art algorithm, we can try to assess the impact of image augmentations, epoch number, model size and the use of pre-trained models when using a small dataset of publicly available images of military materiel. The main goal is to identify Blue and Red forces, ie. Friendly and Foe tanks.

1) The Model: When speaking about DL based object detection one of the leading CNN based models that exist is You Only Look Once (YOLO). Based on the human way of glancing at an image and instantly identifying the objects in the image, their locations and interactions. This allows humans to perform complex tasks with little to no conscious thought. YOLO looks at the whole image and makes predictions based on global context in the image, as well as performing them with a single evaluation, unlike R-CNNs which require numerous evaluations for a single image. [7]

Here we use the latest available YOLOv8 which is an evolution on the previous iterations with greater robustness and enhanced performance, as seen in Figure 1.

Apart from selecting the proper Model, we need to select the size of the model. For the following experiments 3 sizes were selected, in order to explore if and how the number of parameters influence accuracy with small datasets. Here we will use the n, s and m variants. Also, the pre-trained weights that will be used are based on the Common Objects in Context (COCO) dataset, where the attained performance is shown in Table 1.

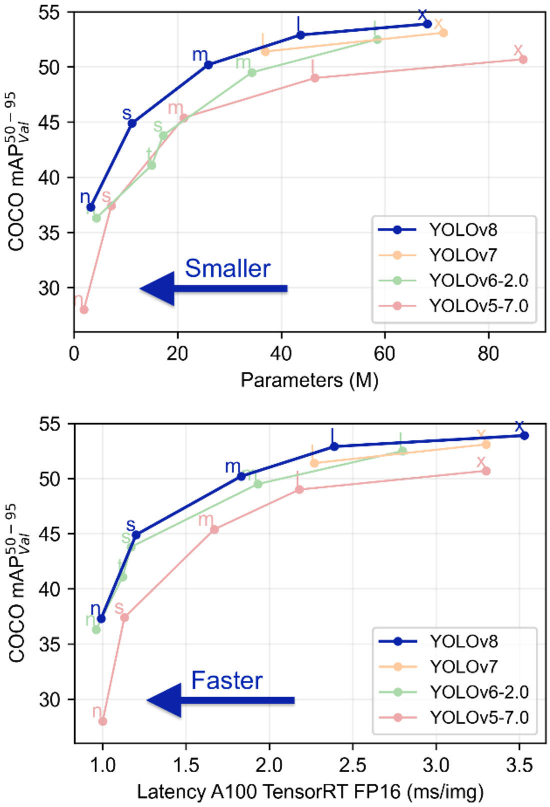

Fig. 1. Comparison of YOLO performance across versions [8]

TABLE 1 COMPARISON OF YOLO MODEL SIZES [8]

Model	Size (pixels)	mAPval 50-95	Params (M)
YOLOv8n	640	37.3	3.2
YOLOv8s	640	44.9	11.2
YOLOv8m	640	50.2	25.9

2) The Dataset: One of the biggest challenges in the Military domain regarding the use of AI techniques is the relative scarcity of available training data, as well as the quality of the data. A large number of the available photos for military materiel is from the official channels from the governments or trade shows where the equipment is shown. These bring the problem too clean a data point, as the photos usually are in perfect lighting conditions, no weather effects at all, as they are taken in hangers or pavilions. These can be used in training, but when a model trained of this data is used in real worlds environments it may struggle with various artefacts, weather conditions, lens distortion and so on. However, even a small dataset is preferable to a non-existent one, so the one used compromises of 1069 images of tanks. These are prepared for You Only Look Once v8 (YOLOv8) by the Roboflow platform and separated in a 70/20/10 split for training, validating and testing respectively. The sample is limiting in one more aspect, as the goal is not only to see the performance of object detection, but also for classification of friendly and foe forces, where both sides can utilize the same or almost the same model of tank. For example, the upgraded Soviet era T-72B3/4 and the Czech version T-72M4. Some examples are shown in Figures 2 and 3.

Fig. 2. Russian T-72B4

Fig. 3. Czech T-72M4

3) The Augmentations: The image augmentation techniques used here are focused mainly on simulating real-world conditions such as distortions and weather effects, as well as the already standard partial images, mosaic distortion, etc. These are implemented through the Albumentations [9] add-on for Ultralytics YOLOv8. The used augmentations include the following:

- ImageCompression
- RandomFog
- RandomRain
- RandomShadow
- RandomSnow
- RandomSunFlare

As well as the standard in-built MixUp augmentations. Examples of the augmented images are shown in Figures 4 and 5.

Fig. 4. Training batch of images with random weather effects

Fig. 5. Training batch of images without random weather effects

B. Results

We will experiment with both re-training of the models and using pre-trained weights on our dataset, as well as the impact of a longer training period (100 epochs vs 200 epochs) and the effect of the augmentations on the accuracy and recall.

1) Re-training versus pre-trained weights: The first variable we will look at is if re-training the models with the very limited available dataset will provide a benefit, or the pretrained models will be able to transfer their domain knowledge to an unknown domain well enough. This has been validated for all three sizes of the Model. See Figures 6, 7 and 8.

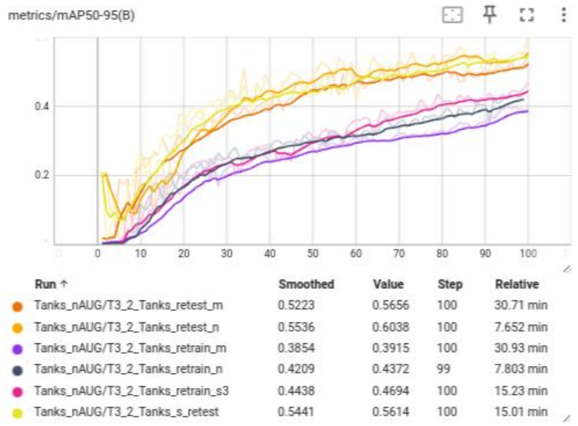

Fig. 6. mAP50 for Re-train vs Pre-trained

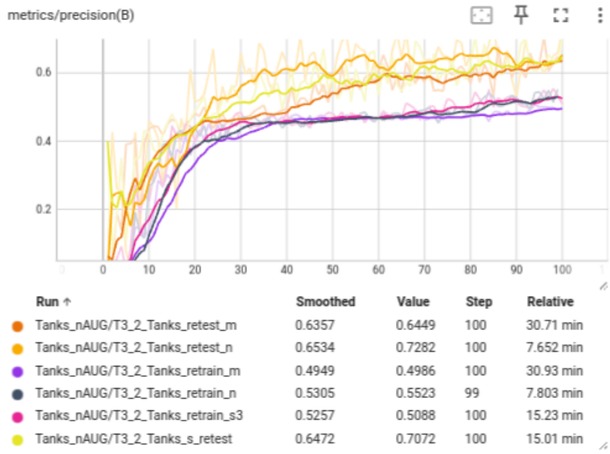

Fig. 7. Precision for Re-train vs Pre-trained

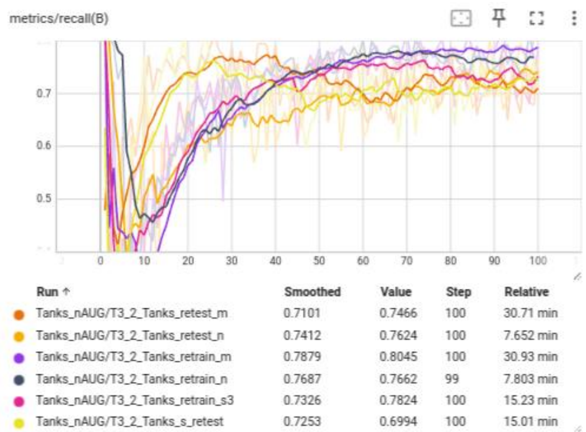

Fig. 8. Recall for Re-train vs Pre-trained

2) Impact of number of epochs: The second variable we will look at is the number of epochs and it's impact, mainly to see if spending more machine time will improve significantly the results. This has been validated for all three sizes of the Model. Also we will look at the time it *took to complete the runs with a single Nvidia RTX 4070 Super card. See Table 2.*

TABLE 2 COMPARISON OF EPOCH NUMBER AND TIME FOR COMPLETION OF RUNS

Model	mAP50	mAP50-95	Precision	Recall	Time
pre-trained N 100	0.7735	0.6038	0.7282	0.7624	7.652 min
pre-trained N 200	0.7437	0.5892	0.6463	0.7817	15.22 min
pre-trained S 100	0.7218	0.5614	0.7072	0.6994	15.01 min
pre-trained S 200	0.7443	0.5917	0.7248	0.6991	30.81 min
pre-trained M 100	0.7236	0.5656	0.6449	0.7466	30.71 min
pre-trained M 200	0.7162	0.5638	0.6429	0.7389	59.84 min

3) Impact of Augmentations: The third variable we will look at is the augmentations in order to see if and how adding them changes the behavior of the training and the results. This has been validated for all three sizes of the Model but only for 100 epochs. See Figures 9 and 10, and Table III.

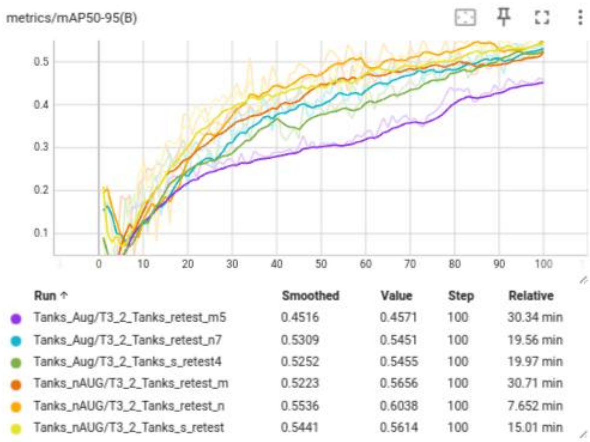

Fig. 9. mAP50-95 for all 3 models with and without augmentations

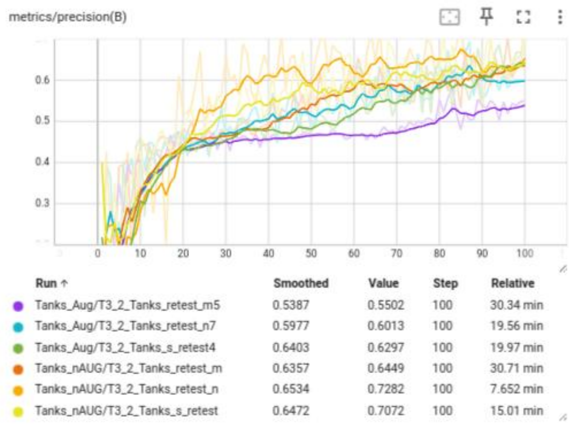

Fig. 10. Precision for all 3 models with and without augmentations

TABLE 3 COMPARISON OF AUGMENTATIONS IMPACT AND SIZE OF MODEL

Model	mAP50	mAP50-95	Precision	Recall	Time
Aug/YOLOv8m	0.6001	0.4571	0.5502	0.7223	30.34 min
Aug/YOLOv8n	0.6996	0.5451	0.6013	0.7778	19.56 min
Aug/YOLOv8s	0.6965	0.5455	0.6297	0.7033	19.97 min
No Aug/YOLOv8m	0.7236	0.5656	0.6449	0.7466	30.71 min
No Aug/YOLOv8n	0.7735	0.6038	0.7282	0.7624	7.652 min
No Aug/YOLOv8s	0.7218	0.5614	0.7072	0.6994	15.01 min

C. Pre-Training vs. Re-Training

The comparison given here shows significant performance disparities. The COCO dataset pre-trained YOLOv8n model achieved 68.2% mAP50 (Mean Average Precision) on test data, compared to 52.4% for the re-trained equivalent – a 30.2% relative improvement. This shows the importance of transfer learning in military AI applications where annotated datasets remain scarce. The pre-trained models' ability to generalize from civilian to military vehicles, this suggests latent feature representations applicable across domains.

D. Model sizes

The size analysis shows counterintuitive results – the YOLOv8s variant outperformed larger (m) and smaller (n) counterparts when using pre-trained weights, achieving 72.1% mAP50 versus 68.9% (m) and 68.2% (n). This suggests that intermediate model complexity optimally balances parameter count and generalization capability for the given dataset size. The inverse relationship between model size and performance in limited-data scenarios highlights the risk of over-parameterization in military AI systems where training data acquisition faces operational constraints.

E. Epochs and Efficiency

Here we see non-linear returns on epoch extension. Doubling epochs from 100 to 200 yielded only 2.8% mAP50 improvement for YOLOv8s, while increasing training time by 87%. This diminishing return suggests practical optimization thresholds exist based on operational urgency and resource availability. By using consumer-grade hardware the study demonstrates feasibility for field-deployable training systems.

Energy efficiency metrics derived from training times and hardware specifications suggest potential for edge-computing implementations. The YOLOv8n model's 2.7-hour training time at 100 epochs versus YOLOv8m's 5.1-hour duration provides commanders with flexibility in model selection based on mission requirements. These findings challenge assumptions about needing specialized hardware for military AI implementations, supporting decentralized training paradigms.

F. Strategies and Trade-Offs

The Albumentations framework implemented six environmental degradation techniques: image compression (simulating transmission artifacts), random fog, rain, shadow, snow, and sun flare. Quantitative analysis reveals a 4.2% decrease in mAP50-95 for YOLOv8s with augmentations enabled, contrasted with a 7.1% recall improvement in foggy conditions. This precision-recall trade-off suggests augmentation strategies must align with mission priorities—surveillance systems prioritizing threat detection may accept lower precision for higher recall, while targeting systems require inverse optimization.

Qualitative analysis of activation maps shows weather augmentations forcing feature focus on structural vehicle components rather than surface textures. This shift improves invariance to environmental conditions but reduces discrimination capability for visually similar variants. The snow augmentation proved particularly detrimental, degrading mAP50 by 9.3% compared to non-augmented baseline. These findings underscore the need for mission-specific augmentation pipelines rather than universal application.

G. Friend-Foe Identification Challenges

The dataset's composition of visually similar adversarial vehicles (Russian T-72B4 vs. Czech T-72M4) tests classification boundaries under operational constraints. Baseline performance without augmentations achieved 84.6% friend-foe accuracy, dropping to 79.1% with weather augmentations enabled. This 5.5% decrease highlights the vulnerability of visual identification systems to environmental degradation, necessitating multi-modal sensor fusion for reliable combat identification.

Error analysis reveals confusion in side-profile views where variant-specific modifications (explosive reactive armor configuration, storage boxes) become occluded. Augmentations exacerbating edge blurring (compression, fog) disproportionately impacted these cases, suggesting terrain-specific augmentation strategies. For urban combat scenarios with frequent partial occlusions, targeted occlusion augmentations during training may improve robustness beyond generic weather effects.

IV. CONCLUSIONS

The integration of AI into military SA systems represents a critical next step in modern defence technology. Through this study we explore the efficacy of image augmentation techniques in enhancing object detection and classification accuracy for friend-foe identification. Key findings show that pre-trained models consistently outperform re-trained counterparts, smaller model architectures demonstrate superior adaptability to limited datasets, and strategic image augmentations improve recall at the expense of precision. Computational efficiency analysis shows diminishing returns with increased training epochs. Weather-based augmentations introduce trade-offs that vary depending on the size of the models used.

These insights show that it is vital to balance the model complexity, training strategies, and augmentation selection to optimize performance where data is scarce, especially in military applications.

A. Situational Awareness Hierarchy

The three-level SA model provides a framework for evaluating AI systems. Current systems are focused on Level 1 (object detection) and Level 2 (contextual classification), with friend-foe identification bridging these levels through semantic understanding. Decision making in the military domain requires not only detection accuracy, but contextual interpretation – the ability to distinguish between Fig. 2 and Fig. 3, this carries different tactical implications than generic vehicle detection.

This hierarchical understanding becomes especially complex when having in mind adversarial deception tactics. The inclusion of near-identical vehicle models from opposing forces in the dataset tests the system's ability to discern subtle visual differences – a capability crucial for preventing fratricide and ensuring target validity. When the focus are these edge cases the study addresses this critical gap in practical military AI applications.

B. Limitations and Future Research Directions

The study's primary limitation stems from its 1,069-image dataset—orders of magnitude smaller than commercial computer vision benchmarks. Performance disparities between pre-trained and re-trained models (15.8% mAP50 difference) directly correlate with this data scarcity. Future work must address military data acquisition challenges through synthetic data generation using game engines or 3D modeling, potentially combined with limited real-world captures.

The dataset's focus on static vehicle imagery neglects motion blur and dynamic engagement scenarios. Incorporating temporal sequences and multi-view captures could improve model robustness to realistic battlefield conditions. Additionally, the absence of civilian vehicles in the dataset limits understanding of false positive rates in complex environments—a critical consideration for urban operations.

Where the study can be taken is with:

1) Augmentation Pipeline Optimization:

Current results suggest blanket augmentation application degrades overall performance, necessitating smarter augmentation strategies. Potential approaches include:

a) Adversarial augmentation: Using reinforcement learning to develop mission-specific augmentation policies

b) Sensor fusion augmentation: Correlating visual degradations with IR/radar sensor characteristics

c) Dynamic scheduling: Gradually introducing harsh augmentations as training progresses

The 6.8% recall improvement under fog conditions demonstrates augmentation value for specific operational environments. Developing weather-prediction-coupled augmentation systems could enable anticipatory model tuning based on mission forecasts. Furthermore, combining geometric transformations (rotation, scaling) with photometric augmentations may better simulate sensor movement in mobile platforms.

2) Multi-Modal Integration Opportunities:

Future systems could enhance visual identification through:

a) RF signature correlation: Matching visual detections with electronic warfare sensor data

b) Acoustic profiling: Cross-validating vehicle engine signatures

c) Contextual awareness: Integrating geographic and mission data for likelihood estimation

The study's visual-only approach achieves militarily useful accuracy (84.6%) but remains below the 95% threshold required for autonomous engagement decisions. Hybrid systems combining visual classification with IFF transponders and battlefield management system inputs could bridge this gap while maintaining redundancy.

C. Closing statement

This comprehensive analysis demonstrates both the potential and limitations of current AI-driven object detection systems in military situational awareness applications. The YOLOv8 architecture, particularly in its s-variant configuration, proves capable of friend-foe identification tasks when leveraging pre-trained weights and strategic augmentation policies. Key implementation insights include the superiority of transfer learning over from-scratch training, optimal model size selection relative to dataset scale, and context-dependent augmentation efficacy.

ACKNOWLEDGMENT

This work was supported by the NSP SD program, which has received funding from the Ministry of Education and Science of the Republic of Bulgaria under the grant agreement no. D01-74/19.05.2022.

REFERENCES

[1] Yibo Sun and Zhe Sun and Weitong Chen, The evolution of object detection methods https://www.sciencedirect.com/science/article/pii/S095219762400616X, 2024.

[2] Iliev, R., A. Genchev. Possibilities for using unmanned aerial vehicles to obtain sensory information for environmental analysis Information & Security: An International Journal 46, no. 2 : 127-140. https://doi.org/10.11610/isij.4609, 2021

[3] Iliev, R., A. Genchev. Generalized net model of the decision making process in the crisis management Journal 'Advanced Studies in Contemporary Mathematics', 31 No 4 pp 401 - 409, http://dx.doi.org/10.17777/ascm2021.31.4.401, 2021

[4] Iliev R., Kochankov M. A generalized net model of command and control system, Vol. 2 Environment. Technology. Resources. Proceedings of the 15th International Scientific and Practical Conference. Volume 2 / Information Technologies, pp. 127 - 131, https://doi.org/10.17770/etr2024vol2.8035, 2024.

[5] Munir, Arslan, Alexander Aved, and Erik Blasch, Situational Awareness: Techniques, Challenges, and Prospects https://doi.org/10.3390/ai3010005, 2022.

[6] Joint Publication 3-09, Joint Fire Support https://web.archive.org/web/20140411125036/http://www.dtic.mil/doctrine/new pubs/jp3 09.pdf, 2010.

[7] Joseph Redmon, Ali Farhadi, YOLOv3: An Incremental Improvement https://pjreddie.com/media/files/papers/YOLOv3.pdf, 2018.

[8] Glenn Jocher and Ayush Chaurasia and Jing Qiu, Ultralytics YOLOv8 https://github.com/ultralytics/ultralytics, 2023.

Buslaev, Alexander, Vladimir I. Iglovikov, Eugene Khvedchenya, Alex Parinov, Mikhail Druzhinin, and Alexandr A. Kalinin, Albumentations: Fast and Flexible Image Augmentations https://doi.org/10.3390/info11020125, 2020.