Journal of Kryvyi Rih National University

Vol. 24, No. 1, 2026

Journal homepage: https://journal-knu.com.ua/en

ISSN 2306-5451
e-ISSN 2519-2736

Article's History: Received: 21.08.2025 Revised: 25.02.2026 Accepted: 28.04.2026 Published: 29.05.2026

UDC 004.93:004.8:355.40

DOI: 10.31721/2306-5451-2026-1-24-107-119

Application of computer vision for target detection in multi-level military reconnaissance systems: Literature review

Denys Berkovskyi *

Postgraduate Student
National Aerospace University "Kharkiv Aviation Institute"
61000, 17 Vadym Manko Str., Kharkiv, Ukraine
https://orcid.org/0009-0004-0625-2839

Alina Artomova

PhD in Technical Sciences, Associate Professor
National Aerospace University "Kharkiv Aviation Institute"
61000, 17 Vadym Manko Str., Kharkiv, Ukraine
https://orcid.org/0000-0002-6761-2066

Abstract. The purpose of the research was to analyse modern computer vision methods for automatic detection and classification of targets in multi-level military intelligence systems, as well as to evaluate the effectiveness of deep learning algorithms in processing intelligence data under various observation conditions. To achieve this goal, the approach integrated a systematic literature review, the modelling of image processing architectures, and a comparative evaluation of neural network-based object detection models. Findings indicated that the deployment of state-of-the-art deep learning architectures specifically convolutional neural networks, YOLO (You Only Look Once) variants, and transformer-based models substantially improves the accuracy and robustness of automated military object detection in imagery acquired from unmanned aerial vehicles and satellite sources. Moreover, combining data from multiple sensors with multimodal processing techniques enhanced target recognition capabilities, particularly in scenarios characterised by low visibility, complex scene composition, or partial object occlusion. A comparative analysis of modern architectures has shown that the YOLO family models provide the best processing speed and are most suitable for real-time application on board unmanned aerial vehicles, while Faster Region-based Convolutional Neural Network demonstrates higher localisation accuracy for complex and small-sized objects, but requires higher computational costs. Convolutional neural network architectures provide a balanced ratio between accuracy and performance, while transformer-based models work more effectively in difficult observation conditions, in particular, with low spatial resolution, noise, partial overlap and masking of objects, but their use is accompanied by increased requirements for computational resources. In general, a compromise has been established between detection accuracy, processing speed and computational complexity of the models, which determines the feasibility of their application depending on the characteristics of the reconnaissance platform and the conditions for performing the task. Information was provided on the development of intelligent systems for automated analysis of intelligence data

Keywords: deep learning; computer vision; object detection; multisensory data fusion; real-time analysis; unmanned aerial vehicle

Suggested Citation:
Berkovskyi, D., & Artomova, A. (2026). Application of computer vision for target detection in multi-level military reconnaissance systems: Literature review. *Journal of Kryvyi Rih National University*, 24(1), 107-119. doi: 10.31721/2306-5451-2026-1-24-107-119.

**Corresponding author (d.s.berkovskyi@khai.edu)*

Introduction

The effectiveness of reconnaissance in current military operations is contingent upon the rapid acquisition, processing, and interpretation of substantial volumes of diverse data originating from multiple sources. The escalating complexity of the battlefield environment, the extensive deployment of unmanned aerial systems, and the expanding volume of visual intelligence have notably impeded the timely identification of military targets. Conventional manual analysis methods are now inadequate to meet the requisite speed and accuracy for decision-making, thereby necessitating the adoption of automated approaches capable of processing reconnaissance data in near real time. Within this domain, computer vision and deep learning technologies have emerged as particularly promising due to their inherent capacity for object detection, classification, and tracking under demanding operational conditions.

Researchers N. Habash *et al*. (2025) noted that in contemporary combat settings, the prompt and accurate acquisition of information regarding the location, movement, and attributes of enemy assets is essential. Intelligence systems serve as the foundation of this process, typically integrating various components such as satellite observation platforms, unmanned aerial vehicles (UAVs), ground-based sensor arrays, and analytical centres responsible for processing and interpreting the collected data. According to W. Zhu & K. Chen (2026), the rapid advancement of digital technologies has led to a substantial increase in reconnaissance data volume, rendering manual analysis more challenging and time-intensive. The researchers emphasised that automated image and video analysis through computer vision techniques presents a viable solution to this problem. Similarly, researchers K. Elgamily *et al*. (2025) concluded that these methods can markedly enhance the efficiency of detecting and identifying military targets, particularly in aerial reconnaissance systems where large quantities of imagery require continuous and rapid assessment. Researchers C. Lee *et al*. (2024) demonstrated that deep learning algorithms have achieved notable success in object detection across diverse imagery types. Furthermore, researchers W. Hua & Q. Chen (2025) reported that contemporary neural network architectures are capable of identifying various military equipment classes even under conditions of image noise, low resolution, or complex backgrounds. These findings highlight the suitability of deep learning approaches for reconnaissance-related applications.

Researchers S. Mei *et al*. (2024) emphasised that the deployment of UAVs has become a key development in modern reconnaissance systems because they facilitate the collection of extensive visual data from operational zones, which can subsequently be analysed automatically by neural networks. In addition, researchers R. Sairam *et al*. (2023) concluded that this integration not only improves target detection accuracy but also reduces the cognitive load on human operators. Similarly, researchers V. Makarichev *et al*. (2025) investigated the influence of image noise on the performance of YOLO-based object detection in UAV imagery. The authors concluded that the application of modern denoising methods significantly improves localisation and classification accuracy, confirming the importance of image pre-processing for reliable computer vision deployment in practical reconnaissance systems. According to researchers Y. Wu & K. Tong (2025), current reconnaissance frameworks increasingly emphasise data fusion from multiple modalities, including optical imagery, infrared sensors, and radar systems. At the same time, researchers J. Leng *et al*. (2024) identified several persistent challenges, including the reliable detection of small objects, recognition of camouflaged equipment, adaptation to changing weather conditions, and limitations imposed by restricted onboard computational resources.

Despite advancements in military reconnaissance utilising computer vision, current research predominantly concentrates on the performance of isolated algorithms or specific use cases. A comprehensive evaluation of computer vision techniques within integrated military reconnaissance systems, which must account for diverse data sources, operational limitations, and varied mission demands, has received less attention. Moreover, comparative analyses of contemporary deep learning models under standardised evaluation metrics are infrequent. These shortcomings underscore the need for a thorough review and comparative assessment of existing methodologies. This study therefore sought to evaluate computer vision techniques for target detection in hierarchical military reconnaissance systems and to identify robust algorithms for the automated interpretation of reconnaissance imagery. To accomplish this purpose, a systematic literature review was undertaken to identify, analyse, and synthesise current research on computer vision applications in military reconnaissance. Relevant publications published primarily between 2021 and 2026 were sourced from the Web of Science, Scopus, IEEE Xplore, SpringerLink, and ScienceDirect databases. The literature search was performed using combinations of keywords including "computer vision", "deep learning", "object detection", "UAV", "remote sensing", "small object detection", "multisensory data fusion", "You Only Look Once (YOLO) models", "convolutional neural networks (CNN)".

Complementing the systematic literature review, comparative evaluation and modelling methodologies were employed to assess the suitability of current computer vision models for reconnaissance tasks. The comparative analysis encompassed metrics such as detection accuracy, precision, recall, F1-score, inference speed, computational complexity, resilience to environmental variations, and applicability for deployment on resource-limited UAV platforms. The quantitative values presented in Table 2 were derived from representative results reported in the reviewed studies and correspond to object detection performance evaluated using standard benchmark datasets and commonly reported metrics, primarily mean Average Precision (mAP). The qualitative ratings used in Table 2 for localisation quality, small object detection capability, and computational complexity were assigned based on a comparative analysis of the reviewed publications. localisation quality and small object detection capability reflect the reported detection performance of the models under challenging conditions, whereas computational complexity was determined according to the computational resources required for model inference, including model size, memory consumption, and processing speed on embedded and edge computing platforms. Conceptual architectural modelling was utilised to generalise identified approaches and evaluate their potential integration into hierarchical military reconnaissance systems. The functional levels of the model were distinguished according to the sequential stages of reconnaissance data processing and the functional responsibilities of system components, including data acquisition, transmission, preprocessing, intelligent analysis, data integration, and decision support, facilitating the development of recommendations for selecting computer vision techniques for automated reconnaissance image interpretation.

Features of Integrating Computer Vision Algorithms into Multi-Level Military Intelligence Systems

Multi-level military reconnaissance systems are designed to collect, process, and distribute information obtained from heterogeneous intelligence assets operating at different observation levels. Their effectiveness depends not only on the capabilities of individual reconnaissance platforms but also on the coordinated interaction between them. The integration of computer vision technologies has become one of the key factors enabling automated processing of large volumes of reconnaissance imagery and supporting timely operational decision-making within hierarchical intelligence architectures. According to researchers X. Bian & J. Ma (2025), multi-level military reconnaissance systems constitute complex information and analytical infrastructures. Similarly, researchers M. Soeleman *et al.* (2023) noted that these systems integrate various platforms for data collection, transmission, processing, and analysis. As shown in the work of researchers M. Vasishta *et al.* (2024), reconnaissance systems generally operate across several observational tiers, including space-based, aerial, and ground components, each characterised by different technical capabilities and operational constraints. Researchers Y. Yu & F. Da (2023) further emphasised that every level provides unique information sources, making effective integration a fundamental requirement for successful intelligence operations.

As noted by researchers Y. Yu & F. Da (2023), the growing volume of reconnaissance data has increased the importance of computer vision algorithms for automated information processing. One of the distinguishing characteristics of hierarchical reconnaissance architectures is the simultaneous use of multiple sensor types, including optical cameras, infrared sensors, radar systems, thermal imagers, and other observation devices. According to researchers L. Ramos & A. Sappa (2025), these sensors generate heterogeneous data with different spatial resolutions, noise characteristics, and acquisition frequencies. Consequently, computer vision algorithms must support multimodal data fusion while ensuring temporal and spatial consistency between information obtained from different reconnaissance platforms. The integration process typically begins with synchronising incoming data streams in both time and geographic coordinates. Satellite imagery usually provides strategic coverage over extensive areas, whereas UAVs supply high-resolution observations of selected regions, and ground-based sensors deliver detailed local information. Before computer vision algorithms are applied, these heterogeneous datasets require preprocessing, including geometric correction, coordinate transformation, temporal synchronisation, and sensor calibration. Such preprocessing enables data acquired from different sources to be represented within a unified reference framework, thereby improving the reliability of subsequent object detection and target tracking.

The location of data processing depends on operational requirements and available computational resources. According to L. Ramos & A. Sappa (2025), UAV platforms frequently perform preliminary image processing onboard to reduce communication latency and bandwidth consumption. Such onboard processing generally includes image enhancement, object localisation, and preliminary target classification using lightweight neural network models. More computationally demanding computer vision tasks are subsequently executed at edge-computing nodes deployed close to operational areas or within centralised intelligence processing centres, where more powerful computing resources allow comprehensive analysis, multisensory fusion, and long-term data storage. This distributed processing architecture balances real-time responsiveness with analytical accuracy. Computer vision techniques are extensively utilised for processing large reconnaissance data streams under near real-time conditions. As noted by L. Ramos & A. Sappa (2025), modern unmanned aerial vehicles continuously transmit high-resolution video requiring immediate analysis to identify potential threats. Researchers J. Leng *et al.* (2024) concluded that conventional image processing methods often struggle under dynamic battlefield conditions, whereas researchers J. Kang *et al.* (2022) demonstrated that deep learning approaches significantly improve both robustness and detection accuracy. Unlike traditional algorithms relying on manually engineered features, CNN automatically learn discriminative visual representations from training data.

J. Kang *et al.* (2022) also emphasised that effective interaction between reconnaissance levels represents one of the principal challenges of hierarchical intelligence systems. Satellite platforms provide extensive geographic coverage but update observations relatively infrequently. In contrast, UAVs deliver high-resolution imagery with greater temporal frequency while remaining constrained by flight endurance and operational range. Ground-based sensors complement these platforms by supplying highly detailed observations within local monitoring zones. Rather than competing, these platforms provide mutually complementary information, making data fusion an essential component of modern reconnaissance architectures. According to researchers Z. Li *et al.* (2022), computer vision algorithms play a central role in integrating information obtained from multiple reconnaissance platforms. Their functions include automated object detection, target classification, localisation, tracking, and feature extraction. The processed results are subsequently transmitted to intelligence management and decision-support systems, where they are combined with operational databases and additional intelligence sources. Decision-support systems utilise these outputs to generate alerts, prioritise detected objects, update the operational picture, and assist analysts and commanders during mission planning and execution.

As shown in the work of Z. Li *et al.* (2022), object detection remains one of the principal computer vision tasks within reconnaissance systems. The researchers classified modern object detection approaches into two-stage and single-stage architectures. Similarly, researchers Z. Cao *et al.* (2023) compared these approaches and concluded that two-stage models, such as Faster R-CNN, generally achieve higher detection accuracy, whereas single-stage models, including the YOLO and SSD families, offer considerably faster inference suitable for real-time aerial reconnaissance. Together, these studies demonstrate that selecting an appropriate detection architecture depends on the balance between computational efficiency and recognition performance required for a particular operational scenario. Researchers have also investigated computer vision performance under adverse observation conditions. Z. Cao *et al.* (2023) reported that low illumination, atmospheric interference, cluttered backgrounds, camouflage, and varying viewing angles significantly complicate target recognition. In agreement with these findings, researchers S. Paheding *et al.* (2024) demonstrated that deep learning models outperform traditional image processing techniques because they learn complex visual representations capable of generalising across challenging environments.

N. Habash *et al.* (2025) further identified small-object detection as one of the major challenges in reconnaissance imagery, particularly when vehicles, artillery systems, or military infrastructure occupy only a limited number of pixels. To address this limitation, the researchers proposed combining multi-scale feature extraction with specialised neural network architectures optimised for detecting small targets. These findings complement the work of Z. Cao *et al.* (2023), who likewise emphasised the importance of adapting detection models to operational battlefield conditions. Another important research direction concerns multimodal data fusion. According to N. Habash *et al.* (2025), combining observations obtained from different sensor modalities provides a more complete representation of the operational environment than relying on a single information source. Similarly, researchers T. Mouatassim *et al.* (2026) concluded that integrating optical, infrared, and radar imagery substantially improves target detection under night time conditions and limited visibility, where individual sensor types may exhibit reduced performance.

T. Mouatassim *et al.* (2026) additionally emphasised the necessity of optimising computer vision algorithms for deployment on resource-constrained reconnaissance platforms. Because UAVs possess limited computational power and energy reserves, researchers have investigated techniques including parameter quantisation, neural network compression, and hardware acceleration to enable efficient onboard inference without significantly degrading detection accuracy. The same researchers also highlighted that successful integration of computer vision requires reliable communication infrastructure. Data acquired by satellites, UAVs, and ground sensors must be securely transmitted through distributed communication networks to edge-processing nodes or centralised intelligence centres. After multisensory fusion and computer vision analysis, processed information is disseminated to command-and-control and decision-support systems, where it contributes to operational awareness, scenario modelling, threat prioritisation, and military planning.

Finally, researchers Y. Wu & K. Tong (2025) demonstrated that decision-support systems represent the final stage of the intelligence processing chain. Rather than replacing human analysts, computer vision technologies provide automatically generated situational information that accelerates decision-making while reducing operator workload. Overall, the reviewed studies consistently indicate that integrating computer vision into multi-level military reconnaissance systems extends beyond the application of object detection algorithms alone. It encompasses coordinated temporal and spatial synchronisation of multisensory data, distributed processing across onboard, edge, and centralised computing environments, multimodal information fusion, and seamless transmission of analytical results to decision-support systems. Continued research in these areas is expected to improve the effectiveness, responsiveness, and operational reliability of modern military intelligence architectures.

Deep Learning Methods for Automatic Detection of Military Objects in Reconnaissance Images

Recent advancements in artificial intelligence have substantially broadened the scope of automated visual data processing and analysis. According to researchers S. Gui *et al.* (2024), recent developments in artificial intelligence have significantly expanded the capabilities of automated image interpretation across various application domains. Similarly, researchers A. Jonnalagadda *et al.* (2024) emphasised that deep learning techniques have become one of the principal drivers of this progress. In military reconnaissance, these techniques enable enhanced automated detection, recognition, and classification of various objects within images and video streams, as also demonstrated by researchers V. Afifah (2026). According to A. Jonnalagadda *et al.* (2024), the visual data processed by modern reconnaissance systems originate from diverse platforms, including Earth observation satellites, unmanned aerial vehicles, stationary or mobile surveillance cameras, and other intelligence assets. These platforms continuously generate extensive volumes of visual information requiring efficient automated analysis. In operational military environments, however, image interpretation is further complicated by dynamic battlefield conditions, varying illumination, atmospheric effects, sensor noise, communication delays, electronic interference, and intentional camouflage. Consequently, computer vision algorithms must maintain high detection accuracy while satisfying strict latency requirements necessary for real-time decision-making.

As noted by A. Jonnalagadda *et al.* (2024), deep learning represents a specialised branch of machine learning that utilises multilayer artificial neural networks capable of identifying complex patterns in large datasets without relying on handcrafted feature descriptors. In contrast, N. Al-IQubaydhi *et al.* (2024) explained that traditional computer vision techniques depended primarily on manually engineered visual features, whereas contemporary neural networks automatically learn hierarchical feature representations during training. The researchers concluded that this capability enables deep learning models to perform effectively even under conditions characterised by substantial variations in object appearance, background complexity, and observation geometry.

According to V. Afifah (2026), CNNs remain among the most widely adopted deep learning architectures for image analysis. These networks are specifically designed for processing two-dimensional image data and employ successive convolutional layers to extract visual features with increasing semantic complexity. Their ability to learn discriminative representations has made CNNs particularly suitable for reconnaissance imagery acquired under heterogeneous operational conditions.

Recent research conducted by V. Makarichev *et al.* (2025) demonstrated that applying image pre-filtering and noise mitigation techniques significantly improves the performance of YOLO-family neural networks for vehicle and human localisation and classification in noisy UAV imagery. These findings suggest that detection performance depends not only on the neural network architecture itself but also on image quality and pre-processing procedures preceding inference. According to researchers M. Nikouei *et al.* (2025), neural network architectures used for automatic military object detection are generally divided into two principal categories: two-stage and single-stage detectors. Similarly, researchers E. Kırac & S. Özbek (2024) concluded that although two-stage detectors generally provide higher localisation accuracy, single-stage models offer considerably faster inference, making them particularly suitable for real-time reconnaissance applications where processing latency is critical. Nevertheless, the choice between these approaches depends on mission requirements, since faster processing often comes at the expense of detection accuracy for complex scenes or small targets.

Transformer-based architectures have recently attracted considerable research interest due to their capability to model long-range contextual relationships within images. According to M. Nikouei *et al.* (2025), this characteristic enables improved object detection in complex environments containing background clutter, camouflage, or significant viewpoint variation. However, compared with conventional CNN, transformer-based models generally require substantially greater computational resources, larger training datasets, and increased memory capacity. Consequently, although these architectures may achieve superior performance under certain conditions, their deployment on resource-constrained edge platforms, including onboard UAV computers, remains challenging. Table 1 presents a comparative analysis of deep learning algorithms, including their mathematical models and key computational principles.

Table 1. Comparative analysis of deep learning algorithms

Algorithm	Model type	Core idea	Strengths	Limitations
YOLO (v5/v8)	Single-stage detector	Direct object detection in one pass	Real-time performance, high speed	Lower accuracy for small objects
Faster R-CNN	Two-stage detector	Region proposal + classification	High detection accuracy	Slower processing speed

Table 1. Continued

Algorithm	Model type	Core idea	Strengths	Limitations
SSD	Single-stage detector	Multi-scale feature maps for detection	Good balance speed/ accuracy	Struggles with small objects
CNN (classification)	Feature extractor	Hierarchical feature learning	Strong feature extraction	Not designed for detection alone
Vision Transformer (ViT)	Transformer-based	Attention mechanism for image patches	High accuracy, global context	High computational cost

Source: developed by the author based on research by E. Kırac & S. Özbek (2024), M. Nikouei *et al.* (2025), A. Jonnalagadda *et al.* (2024) and V. Afifah (2026)

As noted by M. Nikouei *et al.* (2025), detecting small objects remains one of the most difficult tasks in automated reconnaissance image analysis. To address this limitation, researchers have proposed multi-scale feature extraction and feature pyramid architectures capable of preserving fine spatial information across multiple image resolutions, as summarised in Table 2. Nevertheless, detection performance is influenced not only by network architecture but also by the availability of high-quality annotated datasets. Small military targets frequently occupy only a limited number of pixels, while inaccurate or incomplete annotations may substantially reduce recognition accuracy even when advanced detection models are employed.

Table 2. The difficulty of detecting small objects

Detection method	Approach type	Working principle	Detection accuracy (%)	Localisation quality	Small object detection capability	Computational complexity
Single-scale detection	Single-scale	Processes the image at a single resolution	68	Low	Low	Low
Multi-scale detection	Multi-scale	Processes the image at multiple resolutions	81	Medium	Medium	Medium
Feature pyramid network	Hierarchical (FPN)	Uses feature pyramids across different scales	89	High	High	Above medium

Notes: detection accuracy values represent representative object detection performance reported in the reviewed studies and are based primarily on the mean Average Precision (mAP) metric evaluated on benchmark datasets for aerial and remote sensing object detection. The qualitative ratings were assigned through a comparative analysis of the reviewed publications, taking into account the ability of each method to accurately determine object positions, reliably detect small and weakly distinguishable targets, and maintain acceptable performance under different computational resource constraints

Source: developed by the author based on research by M. Soeleman *et al.* (2023), A. Jonnalagadda *et al.* (2024) and M. Niouei *et al.* (2025)

Beyond object detection, contemporary image analysis systems increasingly incorporate semantic and instance segmentation techniques. According to researchers Y. Wu & K. Tong (2025), these methods enable precise delineation of object boundaries and are particularly valuable in cluttered operational environments where overlapping or irregularly shaped objects frequently occur. Compared with conventional bounding-box detection, segmentation approaches provide richer spatial information that may improve subsequent target identification and scene understanding. Multisensory data integration has become another important research direction. According to researchers Z. Li *et al.* (2022), combining optical, infrared, and radar imagery enables more reliable target detection under varying operational conditions. Y. Wu & K. Tong (2025) similarly concluded that multimodal fusion enhances system robustness when individual sensors experience degraded performance due to weather conditions, smoke, poor illumination, or intentional concealment. Tracking detected objects across consecutive video frames constitutes another essential task in reconnaissance applications. According to Z. Li *et al.* (2022), following initial object detection, motion estimation and temporal feature propagation improve tracking continuity for moving targets. Y. Wu & K. Tong (2025) further emphasised that temporal analysis significantly reduces false detections while improving target trajectory estimation during prolonged surveillance missions. Military reconnaissance environments frequently involve deliberate countermeasures intended to reduce detection effectiveness, including camouflage, decoys, electronic interference, and rapidly changing operational conditions.

Another active research area concerns optimising neural network deployment on computationally constrained platforms. According to researchers J. Leng *et al.* (2024), techniques including model pruning, parameter quantisation, and lightweight network architectures are commonly applied to reduce computational complexity. However, the researchers also noted that increasing computational efficiency generally introduces trade-offs with recognition accuracy, requiring an appropriate balance between performance and resource consumption depending on operational requirements. Overall, the reviewed studies consistently demonstrate that deep learning techniques have become integral to automated processing of reconnaissance imagery. At the same time, the comparative analysis indicates that no single neural network architecture is universally optimal for all military reconnaissance scenarios. Instead, selecting an appropriate solution requires balancing detection accuracy, computational efficiency, robustness to adverse observation conditions, latency constraints, and compatibility with resource-limited deployment platforms.

Problems and Prospects for the Development of Intelligent Target Recognition Systems in Military Intelligence

Advancements in information technologies and artificial intelligence have opened new avenues for automating various aspects of military reconnaissance. According to researchers L. Ma *et al.* (2019), a key focus within this domain is the development of intelligent target recognition systems designed to autonomously process extensive volumes of visual data and identify objects of potential operational significance. The researchers also noted that, despite notable progress, the practical implementation of these systems within actual military information infrastructures remains constrained by a range of technical, methodological, and organisational challenges. One significant challenge stem from the complexity of the visual environments encountered during reconnaissance missions. As noted by researchers T. Mouatassim *et al.* (2026) and L. Jin *et al.* (2026), unlike many civilian computer vision applications where imaging conditions are relatively controlled, military reconnaissance images are frequently subject to noise, inconsistent illumination, intricate background textures, and atmospheric disturbances. Similarly, researchers G. Chen *et al.* (2022) emphasised that targets of interest may be partially or fully obscured by natural or artificial elements such as foliage, buildings, or camouflage installations.

According to G. Chen *et al.* (2022), the high variability in the visual appearance of military equipment and infrastructure further complicates automatic object recognition. The researchers concluded that an identical target could present substantially different visual features depending on factors such as viewing angle, lighting, sensor distance, or applied camouflage. Consequently, computer vision algorithms must exhibit robust generalisation capabilities to accurately detect objects despite significant deviations from training examples. Another impediment is the scarcity of labelled training data. According to researchers K. Wang *et al.* (2025), deep learning models typically demand large, annotated datasets comprising numerous examples from diverse object categories. However, in the military context, such datasets are often inaccessible due to security and confidentiality constraints. Therefore, the researchers proposed alternative strategies, including the synthesis of training data and the adoption of transfer learning methodologies. As noted by K. Wang *et al.* (2025), transfer learning entails leveraging models pre-trained on extensive publicly available datasets, such as ImageNet or COCO, subsequently fine-tuning them with smaller, domain-specific datasets containing military-relevant imagery. According to the researchers, this approach can mitigate the volume of labelled data required while maintaining satisfactory model performance.

High processing speed constitutes another crucial requirement for reconnaissance systems. According to researchers C. Lee *et al.* (2024), operational scenarios frequently necessitate real-time or near-real-time analysis. For instance, an unmanned aerial vehicle surveying an area may generate continuous video streams at dozens of frames per second, each demanding prompt analysis to reliably detect potential targets. In accordance with researchers J. Leng *et al.* (2024), both algorithmic efficiency and hardware capabilities must therefore meet stringent computational performance standards. To address these demands, researchers S. Johnson *et al.* (2026) reported that substantial efforts have focused on optimising neural networks by reducing parameter counts, designing compact architectures, and employing specialised accelerators such as graphics processing units or neural processing units. The researchers further noted that edge computing paradigms, in which data processing is partially executed onboard reconnaissance platforms rather than relying solely on centralised systems, have gained increasing prominence.

Reliability and robustness also represent critical attributes for computer vision systems in military applications, where misidentification can entail severe consequences. According to S. Johnson *et al.* (2026), recognition systems must sustain high accuracy and dependability under challenging environmental conditions. Similarly, researchers M. Nikouei *et al.* (2025) concluded that multisensory data fusion is one promising technique for enhancing reliability. By integrating data from multiple sensor modalities, such as combining optical imagery with infrared or radar inputs, the system can offset individual sensor limitations and achieve improved detection performance, especially under adverse weather or low-light scenarios. Another promising research direction is the exploration of self-supervised and semi-supervised learning approaches. According to researchers S. Gui *et al.* (2024), these methods enable training on extensive unlabelled image collections alongside limited annotated datasets, thereby expanding the effective training data pool and potentially enhancing model generalisation when labelled data are restricted.

S. Gui *et al*. (2024) also emphasised that integrating intelligent recognition systems into comprehensive military information architectures presents additional challenges. According to the researchers, outputs from computer vision algorithms must be seamlessly incorporated into command-and-control frameworks and decision-support tools, necessitating the development of specialised interfaces that facilitate operator interaction with detected object information for informed operational decision-making. In the same study, S. Gui *et al*. (2024) further highlighted that the technologies supporting big data processing and distributed computing play a vital role, since modern reconnaissance infrastructures may aggregate data from numerous observation sources concurrently, requiring software and hardware platforms capable of scalable, high-performance processing.

The development of autonomous reconnaissance systems that operate with limited human oversight is also gaining attention. As noted by S. Gui *et al*. (2024), in such systems, computer vision algorithms contribute not only to object detection but also to navigation, route planning, and tactical decision processes. Looking ahead, S. Gui *et al*. (2024) suggested that continued progress in artificial intelligence is expected to further enhance automated reconnaissance data analysis. The researchers identified the design of next-generation neural network architectures, multimodal data processing techniques, and the advancement of explainable AI systems capable of providing transparent and interpretable rationale for their outputs as particularly promising research directions. In summary, computer vision-based intelligent target recognition systems hold substantial promise for augmenting the effectiveness of military reconnaissance operations. While technical challenges persist, ongoing advancements in AI methodologies, computational resources, and data processing algorithms are progressively fostering conditions conducive to the broader adoption of these systems within contemporary military information infrastructures.

Architecture of an Intelligent Intelligence Image Processing System in Multi-Level Intelligence Complexes

The effective application of computer vision techniques in military intelligence largely depends on the proper design of the architecture of an intelligent system for processing reconnaissance data. According to researchers K. Gromada *et al*. (2022), such systems must not only enable automatic detection of objects within images but also integrate information from diverse sources, analyse it, and communicate results to decision support systems. The researchers further noted that contemporary multi-tiered reconnaissance complexes are constructed as intricate information systems comprising several interconnected levels for data collection, transmission, and processing.

According to researchers A. Howard *et al*. (2017), the overall architecture of an intelligent target recognition system in military reconnaissance can be conceptually divided into several primary functional layers: data acquisition, preliminary information processing, intelligent analysis, data integration, and decision support. Similarly, researchers F. Kadhum *et al*. (2026) emphasised that each layer performs specific functions while ensuring interaction with other system components. The initial stage involves collecting data from various observation sources. As noted by researchers V. Makarichev *et al*. (2025), key sources of reconnaissance information include satellite remote sensing systems, unmanned aerial vehicles, ground surveillance cameras, radar stations, and other sensor platforms. According to F. Kadhum *et al*. (2026), each of these sources generates substantial volumes of data, including both static images and high-resolution video streams.

Following data acquisition, preliminary information processing is conducted. According to F. Kadhum *et al*. (2026), the main objective at this stage is to prepare images for subs equent analysis by performing noise filtering, illumination normalisation, correction of geometric distortions, and image alignment. The researchers also noted that images may be transformed into different colour spaces or resised to optimise the performance of deep learning algorithms. Particular emphasis is placed on image georeferencing during this phase. Similarly, researchers R. Beard & T. McLain (2012) concluded that georeferencing represents a critical component of reconnaissance data integration because it allows information from multiple sources to be aligned within a unified coordinate system.

The subsequent stage is the intelligent analysis of images using computer vision algorithms and deep learning methods. According to researchers R. Sairam *et al*. (2023), this layer addresses automatic object detection, classification, and segmentation. The researchers explained that detection algorithms identify the presence and approximate localisation of objects of interest, while classification determines object categories, including military equipment, vehicles, engineering structures, or other infrastructural elements. In more complex scenarios, researchers G. Chen *et al*. (2022) reported that segmentation methods are employed to delineate precise object boundaries. R. Sairam *et al*. (2023) further emphasised that segmentation is particularly valuable when analysing scenes containing overlapping objects or irregular shapes because it provides detailed information regarding scene structure and spatial relationships.

After initial image analysis, results are transferred to the data integration layer. According to G. Chen *et al*. (2022), information obtained from disparate sources is consolidated at this stage. Similarly, R. Sairam *et al*. (2023) demonstrated that satellite image analysis may be combined with data from unmanned aerial vehicles or ground sensor systems to produce a more comprehensive and reliable situational overview. G. Chen *et al*. (2022) identified resolving inconsistencies between heterogeneous information sources as one of the principal challenges of data integration. According to R. Sairam *et al*. (2023), specialised data fusion algorithms evaluate the reliability of incoming observations and generate coherent results. For example, researchers J. Kang *et al*. (2022) explained that when one sensor detects an object while another fails to confirm it, the system must estimate the credibility of each source before assigning an appropriate confidence level to the integrated output.

J. Kang *et al*. (2022) further noted that object tracking modules constitute another essential component of intelligent reconnaissance architectures. Once an object has been detected, the system should continuously monitor its movement over time. According to the researchers, motion analysis algorithms estimate trajectories and predict future positions. Information processed at earlier stages is subsequently conveyed to decision support systems. G. Chen *et al*. (2022) concluded that these systems assist operators by generating notifications regarding newly detected objects, changes in their locations, or suspicious activities. Overall, J. Kang *et al*. (2022) described the architecture of an intelligent image processing system as an interconnected network of components supporting the continuous flow of intelligence data from acquisition to decision-making and adaptive control.

At the initial stage, data acquisition relies on various sensors and sources such as unmanned aerial vehicles, satellites, radar stations, optical cameras, and infrared detectors. Each source delivers raw data characterised by differing spatial resolutions, temporal frequencies, and spectral bands, necessitating careful coordination to facilitate compatibility and effective integration during subsequent processing. Once collected, the data is managed by a transmission and reception module responsible for linking distributed sensors with central processing units. According to J. Kang *et al*. (2022), this module maintains synchronisation across multiple data streams, buffers temporary data when necessary, and ensures secure and reliable communication channels. The researchers emphasised that these functions are essential for preventing data loss or temporal misalignment before downstream processing. The pre-processing module follows, preparing incoming data for further analysis through noise reduction, image stabilisation, normalisation, and georeferencing. As noted by the previously discussed studies, this stage harmonises heterogeneous sensor inputs to support direct data fusion and consistent analysis throughout the system. Subsequently, the intelligent analysis module receives pre-processed data to execute object detection, classification, tracking, multisensory fusion, and situational assessment. According to G. Chen *et al*. (2022), this module employs advanced machine learning methods, including CNNs and transformer architectures, to extract meaningful information from complex reconnaissance imagery. The structured analytical results produced then serve as input for the decision-making subsystem.

Recent research conducted by R. Tsekhmystro *et al*. (2025) suggests that applying pre-processing techniques such as image denoising can significantly improve the effectiveness of deep learning models for object localisation and classification in UAV imagery containing noise. The researchers demonstrated that denoising methods such as DRUNet improve image quality and consequently increase detection accuracy without requiring retraining of existing neural networks. According to G. Chen *et al*. (2022), the decision support module processes analytical outputs to generate actionable intelligence by evaluating operational scenarios, planning responses, and formulating recommendations for operators or command authorities. The researchers emphasised that close integration between the intelligent analysis and decision support modules enables timely and accurate operational decision generation. To facilitate human interpretation and response, researchers L. Jin *et al*. (2026) proposed a visualisation and user interface module that translates analytical outputs into intuitive displays. According to the researchers, these real-time visualisations include detected objects, trajectories, fused sensor information, and situational summaries, thereby bridging automated analysis with operator understanding. A. Rouhi *et al*. (2024) further noted that a feedback control mechanism interconnects the decision support and intelligent analysis components with the data acquisition subsystem. This feedback loop enables dynamic adjustment of sensor configurations, mission parameters, and data collection strategies according to analytical results and operational requirements, thereby improving the adaptability of reconnaissance systems.

Technologies for processing large volumes of data also play a significant role in the operation of modern reconnaissance systems. According to P. Sun *et al*. (2022), the continuously increasing amount of information originating from diverse observation sources requires efficient storage and analysis solutions. Similarly, researchers C. Li *et al*. (2025) concluded that distributed computing systems, cloud platforms, and specialised databases are widely employed to satisfy these computational requirements. L. Jin *et al*. (2026) additionally emphasised that security and reliability remain essential characteristics of military intelligence infrastructures because these systems process highly sensitive operational information. Accordingly, various cryptographic protection mechanisms, access control systems, and redundancy techniques are implemented to ensure resilience against unauthorised access and cyberattacks. Looking forward, L. Jin *et al*. (2026) suggested that future intelligent military reconnaissance architectures will increasingly incorporate artificial intelligence technologies and autonomous capabilities. According to the researchers, next-generation reconnaissance systems may independently analyse operational environments, identify potential threats, and formulate recommendations for subsequent actions with minimal human intervention. Overall, the architecture of an intelligent system for processing reconnaissance imagery constitutes a multi-layered structure comprising interconnected components responsible for data collection, processing, analysis, and integration, as shown in Figure 1.

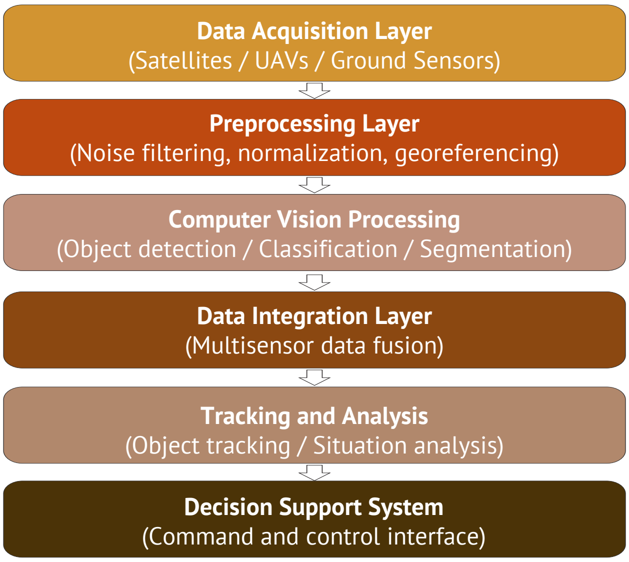

Figure 1. Image processing system architecture
Source: developed by the author based on research by G. Chen *et al.* (2022), P. Sun *et al.* (2022), A. Rouhi *et al.* (2024), C. Li *et al.* (2025) and L. Jin *et al.* (2026)

According to the reviewed studies, the combination of advanced computer vision and deep learning methods with an efficient computational infrastructure substantially enhances the operational effectiveness of contemporary military reconnaissance systems. The reviewed studies indicate that the effectiveness of intelligent reconnaissance architectures depends not only on the performance of individual computer vision algorithms but also on the coordinated interaction of all architectural layers. In practical deployments, several important architectural trade-offs must be considered. Processing data directly on edge platforms, such as UAVs, enables low-latency target detection but is constrained by limited computational resources, whereas centralised or cloud-based processing provides greater analytical capabilities at the cost of increased communication latency. Likewise, highly accurate deep learning models generally require more computational resources, creating a balance between recognition performance and real-time responsiveness. Multimodal data fusion substantially improves situational awareness and detection reliability; however, integrating heterogeneous sensor data increases algorithmic complexity and computational overhead. Finally, although autonomous computer vision systems can significantly reduce operator workload, human involvement in the decision-making loop remains essential for validating critical intelligence outputs and minimising the risks associated with false detections or incorrect classifications. These architectural compromises largely determine the practical effectiveness of modern military reconnaissance systems and should therefore be carefully considered during system design and deployment.

Conclusions

This study explored the application of computer vision and deep learning techniques for the automated detection of military objects within multi-level reconnaissance systems. A review of current scientific literature and existing technological frameworks indicated that advanced image processing algorithms are among the most promising approaches for enhancing the automated analysis of reconnaissance data. The investigation examined the fundamental principles guiding the design of multi-level reconnaissance systems, demonstrating that their operational effectiveness depends significantly on the capability to process large quantities of visual data generated by diverse sensor platforms. In this context, AI-driven computer vision algorithms enable automatic object detection, classification, segmentation, and tracking within video streams, thereby improving processing speed, reducing operator workload, enhancing target detection accuracy, and supporting more efficient decision-making in complex operational scenarios.

Particular emphasis was placed on assessing deep learning approaches for object recognition in reconnaissance imagery. Contemporary neural network architectures, including CNN, single-stage and two-stage detection models, as well as transformer-based frameworks, have exhibited robust performance in computer vision tasks. The analysis demonstrated that no single architecture is universally optimal for all reconnaissance scenarios. Instead, model selection should be based on operational requirements, balancing detection accuracy, inference speed, computational complexity, and suitability for deployment on resource-constrained reconnaissance platforms. Consequently, convolutional, transformer-based, and hybrid approaches should be regarded as complementary rather than competing solutions. The conducted analysis allowed the generalisation of the architectural organisation of multi-level reconnaissance systems. Their operation follows a hierarchical processing pipeline in which reconnaissance data are first collected from heterogeneous sources, including satellites, unmanned aerial vehicles, and ground-based sensors, then synchronised in time and geographic coordinates, pre-processed, analysed using computer vision algorithms, integrated through multisensory data fusion mechanisms, and finally transferred to decision-support systems. Feedback mechanisms subsequently allow analytical results to influence sensor tasking and subsequent data acquisition, forming a closed intelligence processing loop. The effectiveness of this architecture therefore depends not only on the performance of individual deep learning models but also on the efficient interaction between all functional layers.

The review also identified several common architectural principles characterising modern intelligent reconnaissance systems. Distributed processing architectures combining onboard, edge, and centralised computing environments provide the most effective balance between computational efficiency and operational responsiveness. Multisensory data fusion consistently improves situational awareness and target detection reliability by compensating for the limitations of individual sensing modalities. The integration of computer vision algorithms with decision-support systems significantly accelerates intelligence analysis while preserving human supervision during critical operational decisions. Modern deep learning algorithms notably improve object detection accuracy, classification performance, and support the near real-time processing of extensive datasets. This capability is especially critical in contemporary military contexts, where rapid information acquisition and analysis can decisively impact operational outcomes. Nonetheless, several challenges remain and warrant further research. These include difficulties in detecting small objects within high-resolution imagery, the scarcity of specialised training datasets for military applications, ensuring algorithm robustness under diverse observation conditions, and the limitations imposed by computational resources on mobile reconnaissance platforms. Prospective research directions involve devising more effective algorithms for small-object detection, enhancing methods for multisensory data integration, improving adaptive resource allocation between onboard, edge, and centralised processing infrastructures, and progressing toward autonomous intelligent reconnaissance systems capable of executing selected analytical tasks while maintaining reliable human oversight. Additionally, leveraging distributed computing and cloud-based platforms for managing large volumes of reconnaissance data presents a promising avenue for further exploration.

Acknowledgements

None.

Funding

The research has been funded by National Research Foundation of Ukraine (accessed on 24 June 2025) within Project No. 2025.06/0037 "A system for detecting and recognising camouflaged and small objects based on the use of modern computer vision technologies" (2025-2026).

Conflict of Interest

None.

References

[1] Afifah, V. (2026). YOLOv8 for object detection: A comprehensive review of advances, techniques, and applications. *International Journal of Advanced Computational Intelligence*, 2(1), 53-61. doi: 10.71129/ijaci.v2i1.pp53-61.

[2] Al-IQubaydhi, N., Alenezi, A., Alanazi, T., Senyor, A., Alanezi, N., Alotaibi, B., Alotaibi, M., Razaque, A., & Hariri, S. (2024). Deep learning for unmanned aerial vehicles detection: A review. *Computer Science Review*, 51, article number 100614. doi: 10.1016/j.cosrev.2023.100614.

[3] Beard, R.W., & McLain, T.W. (2012). *Small unmanned aircraft theory and practice*. Princeton & Oxford: Princeton University Press.

[4] Bian, X., & Ma, J. (2025). Deep learning based real time detection and localisation for targets in UAV remote sensing images. *International Journal of Advance in Applied Science Research*, 4(9), 8-13.

[5] Cao, Z., Kooistra, L., Wang, W., Guo, L., & Valente, J. (2023). Real time object detection based on UAV remote sensing: A systematic literature review. *Drones, 7*, article number 620. doi: 10.3390/drones7100620.

[6] Chen, G., Liu, S.-J., Sun, Y.-J., Ji, G.-P., Wu, Y.-F., & Zhou, T. (2022). Camouflaged object detection via context-aware cross-level fusion. *ArXiv*. doi: 10.48550/arXiv.2207.13362.

[7] Elgamily, K.M., Mohamed, M.A., Abou Taleb, A.M., & Ata, M.M. (2025). Enhanced object detection in remote sensing images by applying metaheuristic and hybrid metaheuristic optimisers to YOLOv7 and YOLOv8. *Scientific Reports*, 15, article number 7226. doi: 10.1038/s41598-025-89124-8.

[8] Gromada, K., Siemiątkowska, B., Stecz, W., Płochocki, K., & Woźniak, K. (2022). Real-time object detection and classification by UAV equipped with SAR. *Sensors (Basel, Switzerland)*, 22(5), article number 2068. doi: 10.3390/s22052068.

[9] Gui, S., Song, S., Qin, R., & Tang, Y. (2024). Remote sensing object detection in the deep learning era – a review. *Remote Sensing*, 16(2), article number 327. doi: 10.3390/rs16020327.

[10] Habash, N., *et al*. (2025). Recent real-time aerial object detection approaches, performance, optimisation, and efficient design trends for onboard performance: A survey. *Sensors (Basel, Switzerland)*, 25(24), article number 7563. doi: 10.3390/s25247563.

[11] Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. *ArXiv*. doi: 10.48550/arXiv.1704.04861.

[12] Hua, W., & Chen, Q. (2025). A survey of small object detection based on deep learning in aerial images. *Artificial Intelligence Review*, 58, article number 162. doi: 10.1007/s10462-025-11150-9.

[13] Jin, L., Wang, R., & Huang, B. (2026). A challenge-driven survey on UAV-based target tracking. *Computers, Materials & Continua*, 88(1), article number 2. doi: 10.32604/cmc.2026.080050.

[14] Johnson, S.L., Moore, H.D., & Porter, M.D. (2026). Toward resilient multi-modal drone detection in cluttered environments: A systems survey of EO/IR, radar, acoustic, LiDAR and RF modalities. *International Journal of Critical Infrastructure Protection*, 54, article number 100870. doi: 10.1016/j.ijcip.2026.100870.

[15] Jonnalagadda, A.V., & Hashim, H.A. (2024). SegNet: Segmented deep learning CNN approach for wildfire detection in UAV imagery. *Remote Sensing Applications: Society and Environment*, 34, article number 101181. doi: 10.1016/j.rsase.2024.101181.

[16] Kadhum, F.M., Abdulhameed, A.A., & Saud, J.H. (2026). Vision-based UAV detection methods using deep learning: A review. *Journal of Al-Qadisiyah for Computer Science and Mathematics*, 18(1), 281-301. doi: 10.29304/jqcsm.2026.18.12504.

[17] Kang, J., Tariq, S., Oh, H., & Woo, S.S. (2022). A survey of deep learning-based object detection methods and datasets for overhead imagery. *IEEE Access*. doi: 10.1109/ACCESS.2022.3149052.

[18] Kırac, E., & Özbek, S. (2024). Deep learning based object detection with UAV equipped with embedded system. *Journal of Aviation*, 8(1), 15-25. doi: 10.30518/jav.1356997.

[19] Lee, C., Son, J., Shon, H., Jeon, Y., & Kim, J. (2024). FRED: Towards a full rotation-equivariance in aerial image object detection. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(4), 2883-2891. doi: 10.1609/aaai.v38i4.28069.

[20] Leng, J., Ye, Y., Mo, M., Gao, C., Gan, J., Xiao, B., & Gao, X. (2024). Recent advances for aerial object detection: A survey. *ACM Computing Surveys*, 56(12), article number 296. doi: 10.1145/3664598.

[21] Li, C., Zhao, R., Wang, Z., Xu, H., & Zhu, X. (2025). RemDet: Rethinking efficient model design for UAV object detection. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(5), 4643-4651. doi: 10.1609/aaai.v39i5.32490.

[22] Li, Z., Wang, Y., Zhang, N., Zhang, Y., Zhao, Z., Xu, D., Ben, G., & Gao, Y. (2022). Deep learning-based object detection techniques for remote sensing images: A survey. *Remote Sensing*, 14(10), article number 2385. doi: 10.3390/rs14102385.

[23] Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., & Johnson, B.A. (2019). Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 152, 166-177. doi: 10.1016/j.isprsjprs.2019.04.015.

[24] Makarichev, V., Tsekhmystro, R., Lukin, V., & Krytskyi, D. (2025). Performance improvement of vehicle and human localisation and classification by YOLO family networks in noisy UAV images. *Information*, 16(12), article number 1087. doi: 10.3390/info16121087.

[25] Mei, S., Lian, J., Wang, X., Su, Y., Ma, M., & Chau, L.-P. (2024). A comprehensive study on the robustness of deep learning-based image classification and object detection in remote sensing: Surveying and benchmarking. *Journal of Remote Sensing*, 4, article number 0219. doi: 10.34133/remotesensing.0219.

[26] Mouatassim, T., Airaj, M., & El Guarmah, E.M. (2026). Safeguarding the skies: The rise of machine learning approaches for cyberattack detection in unmanned aerial systems. *Cybersecurity* 9, article number 62. doi: 10.1186/s42400-025-00466-2.

[27] Nikouei, M., Baroutian, B., Nabavi, S., Taraghi, F., Aghaei, A., Sajedi, A., & Ebrahimi Moghaddam, M. (2025). Small object detection: A comprehensive survey on challenges, techniques and real-world applications. *Intelligent Systems with Applications*, 27, article number 200561. doi: 10.1016/j.iswa.2025.200561.

[28] Paheding, S., Saleem, A., Siddiqui, M.F.H., Rawashdeh, N., Essa, A., & Reyes, A.A. (2024). Advancing horizons in remote sensing: A comprehensive survey of deep learning models and applications in image classification and beyond. *Neural Computing and Applications*, 36, 16727-16767. doi: 10.1007/s00521-024-10165-7.

[29] Ramos, L.T., & Sappa, A.D. (2025). A decade of You Only Look Once (YOLO) for object detection: A review. *ArXiv*. doi: 10.48550/arXiv.2504.18586.

[30] Rouhi, A., Arezoomandan, S., Kapoor, R., Klohoker, J., Patal, S., Shah, P., Umare, H., & Han, D. (2024). An overview of deep learning in UAV perception. In *2024 IEEE international conference on consumer electronics (ICCE)*. Las Vegas: IEEE. doi: 10.1109/ICCE59016.2024.10444237.

[31] Sairam, R.V.C., Keswani, M., Sinha, U., Shah, N., & Balasubramanian, V.N. (2023). ARUBA: An architecture-agnostic balanced loss for aerial object detection. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision (WACV 2023)* (pp. 3719-3728). Waikoloa, USA: IEEE Computer Society.

[32] Soeleman, M.A., Supriyanto, C., & Purwanto. (2023). Deep learning model for unmanned aerial vehicle based object detection on thermal images. *Revue d'Intelligence Artificielle*, 37(6), 1441-1447. doi: 10.18280/ria.370608.

[33] Sun, P., Liu, T., Chen, X., Zhang, S., Zhao, Y., & Wei, S. (2022). Multi-source aggregation transformer for concealed object detection in millimeter-wave images. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(9), 6148-6159. doi: 10.1109/TCSVT.2022.3161815.

[34] Tsekhmystro, R., Lukin, V., & Krytskyi, D. (2025). UAV image denoising and its impact on performance of object localisation and classification in UAV images. *Computation*, 13(10), article number 234. doi: 10.3390/computation13100234.
[35] Vasishta, M.S.S.V., Amireddy, A.T.R., & Shrivastava, P. (2024). Small object detection for UAVs using deep learning models on edge computing: A comparative analysis. In *2024 5th international conference on circuits, control, communication and computing* (pp. 1-7). Bangalore: IEEE.
[36] Wang, K., Wang, Z., Li, Z., Su, A., Teng, X., Pan, E., Liu, M., & Yu, Q. (2025). Oriented object detection in optical remote sensing images using deep learning: A survey. *Artificial Intelligence Review*, 58, article number 350. doi: 10.1007/s10462-025-11256-0.
[37] Wu, Y., & Tong, K. (2025). Research advances on deep learning based small object detection in UAV aerial images. *Acta Aeronautica et Astronautica Sinica*, 46(3). doi: 10.7527/S1000-6893.2024.30848.
[38] Yu, Y., & Da, F. (2023). Phase-shifting coder: Predicting accurate orientation in oriented object detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR 2023)* (pp. 13354-13363). Vancouver: IEEE Computer Society.
[39] Zhu, W., & Chen, K. (2026). Real time object detection for unmanned aerial vehicles based on vision transformer and edge computing. *Scientific Reports*, 16, article number 6814. doi: 10.1038/s41598-026-37938-5.

Застосування комп'ютерного зору для виявлення цілей у багаторівневих системах військової розвідки: огляд літератури

Денис Берковський

Аспірант
Національний аерокосмічний університет «Харківський авіаційний інститут»
61000, вул. Вадима Манька, 17, м. Харків, Україна
https://orcid.org/0009-0004-0625-2839

Аліна Артьомова

Кандидат технічних наук, доцент
Національний аерокосмічний університет «Харківський авіаційний інститут»
61000, вул. Вадима Манька, 17, м. Харків, Україна
https://orcid.org/0000-0002-6761-2066

Анотація. Метою дослідження був аналіз сучасних методів комп'ютерного зору для автоматичного виявлення та класифікації цілей у багаторівневих системах військової розвідки, а також оцінювання ефективності алгоритмів глибокого навчання в обробці розвідувальних даних за різних умов спостереження. Для досягнення цієї цілі підхід інтегрує систематичний огляд літератури, моделювання архітектур обробки зображень та порівняльну оцінку моделей виявлення об'єктів на основі нейронних мереж. Результати дослідження показали, що розгортання найсучасніших архітектур глибокого навчання, зокрема згорткових нейронних мереж, варіантів YOLO та моделей на основі трансформаторів, суттєво покращує точність та надійність автоматичного виявлення військових об'єктів на зображеннях, отриманих з безпілотних літальних апаратів та супутникових джерел. Крім того, поєднання даних з кількох датчиків з мультимодальними методами обробки покращує можливості розпізнавання цілей, особливо в сценаріях, що характеризуються низькою видимістю, складним складом сцени або частковим перекриттям об'єкта. Порівняльний аналіз сучасних архітектур показав, що моделі сімейства YOLO забезпечують найкращу швидкість обробки та є найбільш придатними для застосування в режимі реального часу на борту безпілотного літального апарату, тоді як Faster R-CNN демонструє вищу точність локалізації складних і малорозмірних об'єктів, але потребує більших обчислювальних витрат. Архітектури згорткових нейронних мереж забезпечують збалансоване співвідношення між точністю та продуктивністю, тоді як моделі на основі трансформерів ефективніше працюють у складних умовах спостереження, зокрема за низької просторової роздільної здатності, наявності шуму, часткового перекриття та маскування об'єктів, однак їх використання супроводжується підвищеними вимогами до обчислювальних ресурсів. Загалом встановлено компроміс між точністю виявлення, швидкістю обробки та обчислювальною складністю моделей, який визначає доцільність їх застосування залежно від характеристик розвідувальної платформи та умов виконання завдання. Надано інформацію про розробку інтелектуальних систем для автоматизованого аналізу даних розвідки

Ключові слова: глибоке навчання; виявлення об'єктів; мультисенсорне об'єднання даних; аналіз у реальному часі; безпілотний літальний апарат