Ukrainian first-person-view (FPV) drones are shown here in a 20 December 2023 photo. Ukraine has produced more than fifty thousand FPV drones. (Photo courtesy of the Ministry of Strategic Industries of Ukraine)

Transforming the Multidomain Battlefield with AI

Object Detection, Predictive Analysis, and Autonomous Systems

Maj. Jim Gallagher, U.S. Army
Dr. Edward J. Oughton

The proliferation and rapid advancement of artificial intelligence (AI) is reshaping the conduct of offensive and defensive operations in multidomain operations (MDO) at unprecedented speed. AI advancements provide myriad new capabilities to the warfighter that were once only imagined as science fiction. AI is accelerating data collection, processing, analysis, and exploitation accuracy at machine speed, thus shortening the OODA (observe, orient, decide, act) loop cycle for commanders. AI is also augmenting processes that were previously done primarily by humans. For example, AI can detect objects of interest across multiple unmanned aircraft system (UAS) footage from various sensor types. Large language models (LLM) can also synthesize big data from different platforms, such as combining imagery, social media posts, and intelligence reports to provide a comprehensive overview of the operational environment (OE) to commanders on demand. AI can also fully automate intelligence, surveillance, and reconnaissance (ISR) platforms and weapon systems. Despite these advancements, AI also brings myriad technical, ethical, and legal challenges when implemented in MDO. This article will discuss these challenges and provide recommendations for the way ahead.

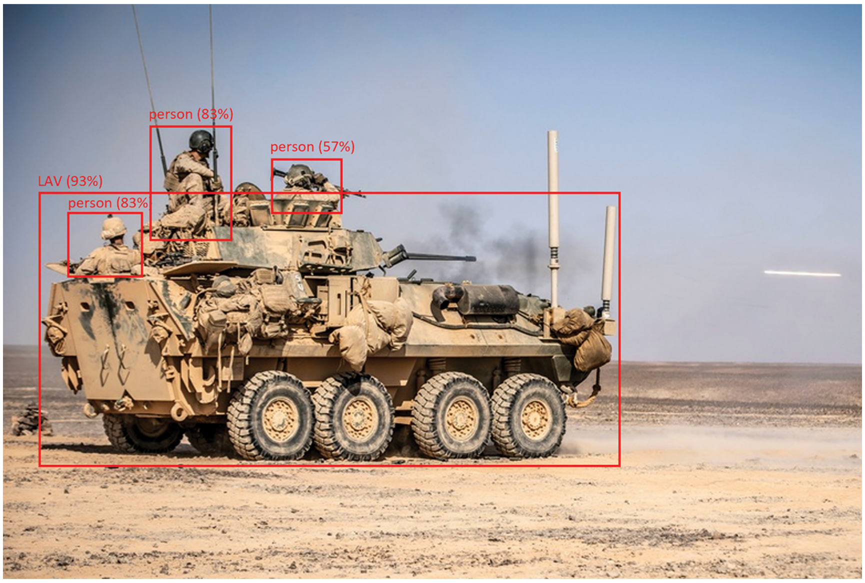
Object Detection in Multidomain Operations

AI is a broad field in computer science focused on creating systems capable of performing tasks that typically require human intelligence. These tasks include decision-making, problem-solving, understanding language, and recognizing patterns or objects.[1] AI models aim to mimic human cognitive functions, which are achieved by applying data to convolutional neural networks (CNN).[2] Data quality and CNN optimization are the two most significant factors for advancing AI models. Machine learning (ML) is a subset of AI that performs specific repetitive functions. Furthermore, ML systems can improve their performance through iterative training cycles, called epochs, without human intervention by adjusting their neural connections within the CNN based on the input data. The most widely used ML application in MDO today is object detection.[3] Object detection uses computer vision to detect objects of interest in images and video footage.[4] Figure 1 is an example of object detection. Within the field of AI, object detection models are the easiest to train and deploy, making them optimal for MDO. Project Maven is an example of an ongoing U.S. military program that applies object detection algorithms to ISR-derived footage and images.[5] However, the challenge in building object detection models, as in any AI model, lies in the need for quality data to pass through the CNN to train the model.

Collecting quality data is a challenging process necessary to train dependable ML models. The performance of object detection models directly depends on the quality of the data on which they are trained.[6] Collecting data during MDO will be one of the biggest hurdles in deploying object detection models trained on new enemy equipment and tactics, techniques, and procedures (TTP) in the OE. The enemy will likely employ techniques to deceive object detection models in MDO. In a high operational tempo fight, collecting, curating, and sharing good data to retrain ML models on emerging enemy equipment and TTPs will be difficult. Data diversity of enemy equipment at various angles and illumination conditions is crucial to achieving resilient object detection models. Furthermore, air-based object detection models are more challenging to train than ground-based models due to decreased edge contrast, thermal crossover (when the temperature of an object

Maj. Jim Gallagher, U.S. Army, is a military intelligence officer serving as the executive officer of the 297th Military Intelligence Battalion (513th Military Intelligence Brigade). His passion for artificial intelligence and unmanned aircraft began during his deployment to Iraq during Operation Inherent Resolve. He holds a BA from Mercyhurst University, an MA from American Military University, an MS from George Mason University, and an MA from the Command and General Staff College; and is a PhD student at George Mason University.

Dr. Edward J. Oughton is an assistant professor in the College of Science at George Mason University, Fairfax, Virginia, developing open-source research software to analyze digital infrastructure deployment strategies. He received the MPhil and PhD degrees from Clare College at the University of Cambridge in 2010 and 2015, respectively. He later held research positions at both Cambridge and Oxford.



A ground-based object detection model identifies objects of interest in this photograph. (Figure by Maj. Jim Gallagher)

Figure 1. Ground-Based Object Detection Model

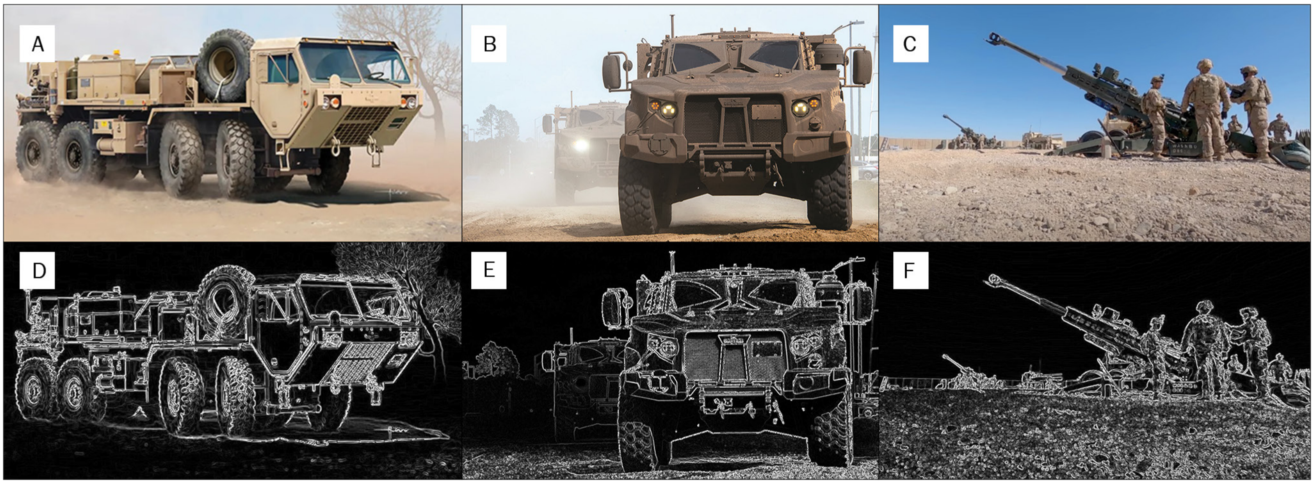
is similar to its background), and image deterioration.[7] Once a dataset of enemy equipment is curated, the unit can build object detection models, which can be tailored to the mission's requirements. For example, in large-scale combat operations, a unit can deploy a custom object detection model trained on a dataset of its high-value targets.

After model training, object detection models can be deployed on cell-phone-sized computers called edge devices. These edge devices are inexpensive and can be easily connected to unmanned vehicle ground stations, providing existing unmanned platforms with off-board AI capabilities. The primary benefit of deploying object detection algorithms in multidomain offensive and defensive operations is their ability to rapidly and consistently process large quantities of video and image footage with minimal human involvement. For example, a division analysis and control element can analyze multiple ISR video feeds with a series of edge devices with custom-trained object detection models. However, the most significant challenge with object detection in MDO is updating hundreds or thousands of edge devices with the latest model trained in new enemy equipment and TTPs. This also includes how to transmit datasets comprised of images and labels to units across the OE. Transmitting these large datasets requires significant bandwidth and time, which will be highly unlikely to execute when facing a near-peer adversary. Currently, there is no workable solution to solving this issue.

Object detection models can also be trained in multiple spectrums, ranging from near infrared (NIR) to long-wave infrared (LWIR), also known as thermal.[8] Choosing the appropriate sensor type to conduct object detection from a ground-based or air-based reconnaissance system is essential for ML applications. Since edge detection is the foundation for computer vision models, choosing the appropriate sensor for ISR object detection performance is critical.[9] Parts *D*/*E*/*F* of figure 2 show how computer vision models "see" an image. Object detection algorithms train themselves to detect the unique external and internal edges of the object to accurately detect and identify objects of interest.

LWIR sensors are optimal for object detection when conducting operations during periods of limited visibility. Figure 3 is an example of an LWIR object detection model identifying a person while an NIR sensor failed to identify the same person. A separate machine learning model on the same edge device can also fuse multiple sensor types to create more resilient object detection models that can continue to perform in complex illumination conditions.[10]

For example, during hours of begin morning nautical twilight (BMNT, one hour before sunrise) and end evening nautical twilight (EENT, one hour after sunset), the visible light camera and the NIR camera from a small UAS can be fused to create compound edges of the target object, thus increasing performance. Previous research has found that employing separate machine-learning models that conduct adaptive sensor fusion increases object detection performance.[11] Figure 4*B* is an example of an image fused with a coaligned RGB and LWIR sensor. A shortfall in applying this method to existing UAS is that most military UAS carry NIR sensors, which are less effective than LWIR for object detection applications.

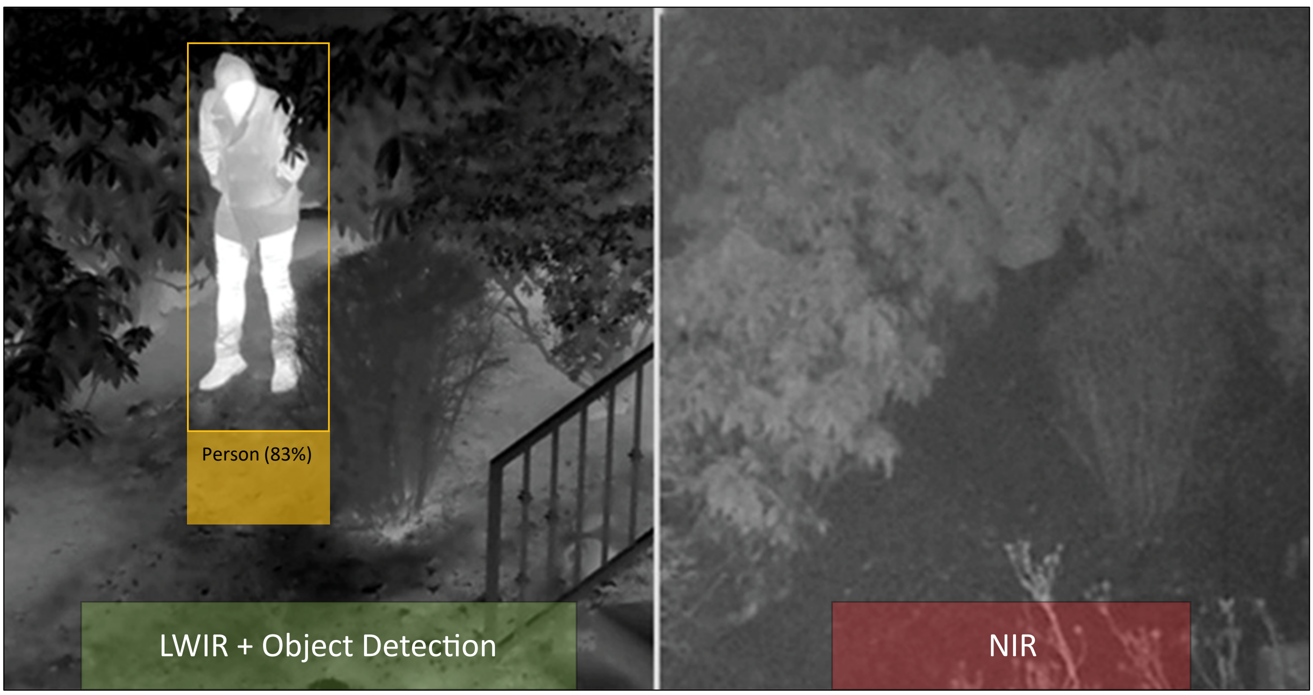


These examples show how a computer vision model perceives images. Edges are highlighted to help with model training and detection. Deteriorating these edges with camouflage is key to defeating adversarial object detection models. (Figure by Maj. Jim Gallagher)

Figure 2. Computer Vision Model Perceiving Images

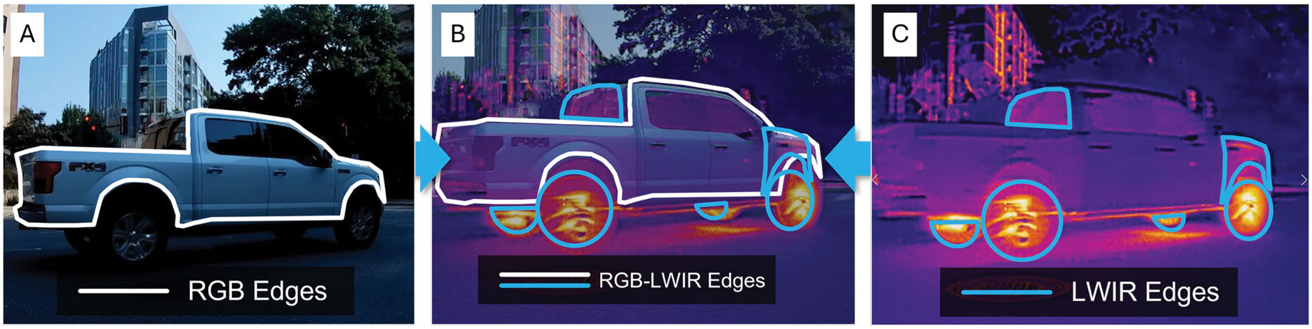
Integrating Object Detection with Targeting

When deploying object detection models in both offensive and defensive operations, detections of enemy equipment derived from ground and air-based sensors will be automatically directed to the strike cell, where a soldier can confirm the target. After the detection is confirmed as hostile, the target and its correlating metadata will begin moving through the D3A (decide, detect, deliver, assess) targeting cycle. Employing object detection models to aid in identifying enemy targets will accelerate the D3A targeting cycle, provide



A drone-based long-wave infrared (LWIR) sensor successfully detects a person (*left image*), while a near infrared (NIR) sensor in the same setting fails to detect the person. (Photo courtesy of Teledyne FLIR)

Figure 3. Drone-Based LWIR/NIR Sensors Detection



Fusing multiple sensors together creates additional and redundant edges, thereby increasing model performance. (Figure by Maj. Jim Gallagher)

Figure 4. Fusing Sensors Together

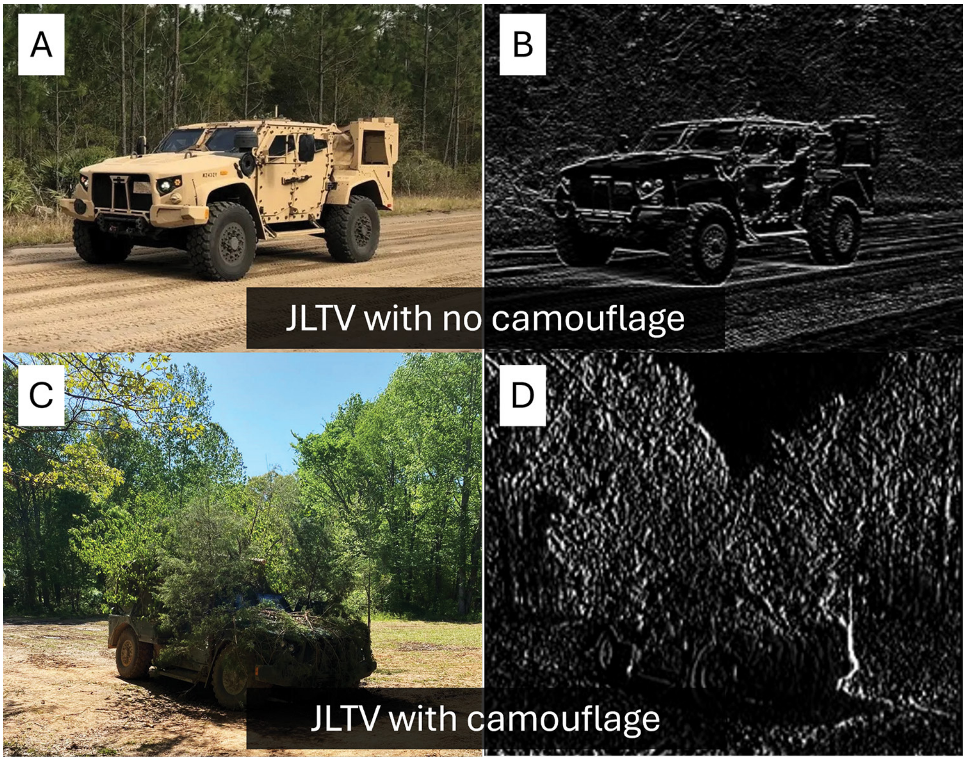
consistent detections, and significantly reduce the number of soldiers required to analyze ISR footage.

A primary vulnerability of object detection models in targeting is that the enemy can degrade or defeat the algorithm with various methods. For example, object detection models employing visible light sensors can be defeated by breaking up the target object's edges with camouflage. Figures 5C and 5D are examples of a Joint Light Tactical Vehicle camouflaged with foliage to break up its edges, resulting in no positive detection from the object detection algorithm. If a soldier is not watching the ISR feed, this would be a missed target. Furthermore, placing a camouflage net over equipment without using spreaders to break up the object's edges is insufficient to defeat object detection models. Figure 6 illustrates a M119 howitzer with protective covers. The covered M119 howitzers were still detected by a ground-based object detection model with a confidence score of 41 percent when covered due to the identifiable silhouette of the howitzer. Foliage and other materials (as shown in figure 5C) must be used to deteriorate an object's edges, thereby decreasing the performance of the object detection model. Because of the limitations and ease of defeating object detection models using visible light sensors, NIR and LWIR sensors are better suited for detecting targets.[12] However, thermal signatures can be masked or completely covered to defeat computer vision models.

Object Detection for Autonomous Munitions

Another AI application that can enhance lethality in multidomain offensive and defensive operations is combining object detection with object tracking to guide munitions into targets autonomously. Munitions can be guided in the air, ground, and sea domains with autonomous vehicles possessing an onboard edge device connected to the vehicle's flight controller. With edge devices costing as low as $35, deploying low-cost unmanned systems with object tracking models is cost-effective and efficient. For example, an unmanned ground vehicle (UGV) can be deployed with an object detection and tracking model to detect and drive ordinance into an enemy target. Figure 7 shows a UGV with an onboard object detection and tracking model that can direct and detonate a low-cost 3D-printed shape charge against an armored vehicle. This UGV was built as part of a project at the Command and General Staff College.

Attaching infrared and visible light sensors to unmanned systems with onboard object detection and tracking will also increase model performance and targeting effects regardless of illumination conditions.[13] The Ukrainian army is already deploying semiautonomous multirotor drones with onboard edge devices to fly munitions into Russian tanks autonomously (figure 8).[14] Unmanned underwater vehicles (UUV) can also be deployed to autonomously drive ordnance into enemy vessels. Additionally, combining object detection models with unmanned systems makes them less susceptible to GPS jamming since GPS is not required for object detection and object tracking. However, a primary risk to utilizing fully autonomous munitions in MDO is the risk of fratricide. Although the possibility of fratricide will be lower in a slow-moving unmanned system such as a UGV or UUV, in a fast-moving UAV, there is an increased risk of fratricide if the object detection



Ground and air sensors with object detection can easily identify a Joint Light Tactical Vehicle with no camouflage (*images A and B*). Conversely, images without detectable edges (*C and D*) are difficult for an object detection model to detect. (Figure by Maj. Jim Gallagher)

Figure 5*A*/*B*. Ground and Air Sensors with Object Detection

model has false positives and detects a friendly system as an enemy. Until neural networks advance further to increase detection confidence to a reasonable level, the concept of deploying a fully autonomous drone swarm with friendly forces nearby is still a distant future thought.

Using Deep Learning for Predictive Analysis

Before discussing how deep learning AI models can be used in MDO, it is necessary to discuss how simple machine learning algorithms must first be used for data processing. Existing ML algorithms, such as speech-to-text converters, text readers, and optical character recognition, can make sense of big data coming into a command post. Machine learning models can process data from upper and lower tactical internet systems to make the data machine-readable for deep-learning models, facilitating rapid situational awareness for the commander.

Covered M119 howitzers are detected by a ground-based object detection model due to identifiable edges. (Figure by Maj. Jim Gallagher)

Figure 6. Detection via a Ground-Based Object Detection Model

This unmanned ground vehicle built at the Command and General Staff College is equipped an onboard object detection and object tracking model that can autonomously aim and initiate the 3D-printed copper shape charge at an armored vehicle. (Photo by Maj. Jim Gallagher)

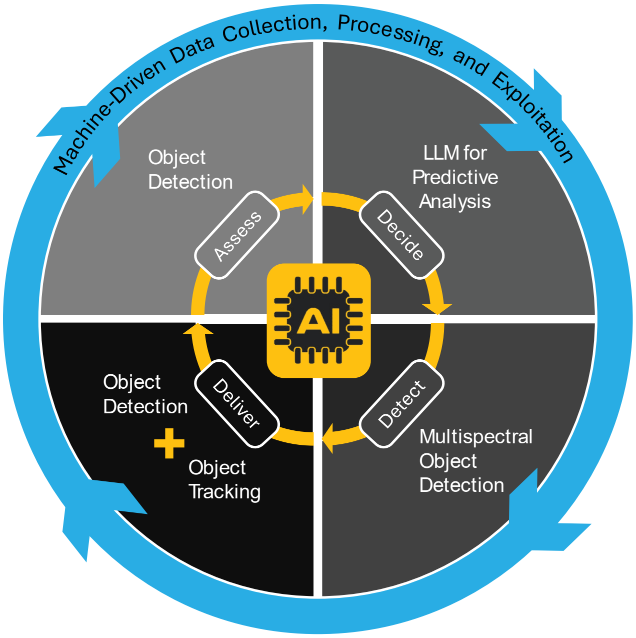
Figure 7. Unmanned Ground Vehicle

After data processing, deep learning algorithms, such as LLMs, can synthesize the data to provide commanders with predictive analysis. This analysis is based on the processed data derived from object detection models on unmanned systems, friendly reporting, and intelligence. The major drawback to LLMs is that they are computationally intensive and difficult to train.[15] Training new LLMs in a field environment is difficult because it requires a large quantity of data, time, and computational resources. Therefore, LLMs should be trained in a garrison environment with ample time and processing power to build the model.

Once the LLM, data processing, and object detection models are built, the models and data pipeline can finally be assembled to conduct AI-driven targeting. Figure 9 visualizes how all the previously discussed AI/ML algorithms would work using D3A as the framework. During the decide portion of D3A, the LLM provides an initial assessment to the commander on the most likely enemy course of action based on known enemy composition and disposition, which is derived from intelligence reports, friendly reporting, and object detection results from unmanned air, ground, and sea systems. Since the LLM is trained on enemy doctrine and has access to OE terrain data, it will provide predictive analysis on the

Ukrainian fighters zip tie a munition to the frame of a low-cost drone. (Photo courtesy of the Ukrainian defense industry via the Ministry of Defense)

Figure 8. Low-Cost Ukrainian Drone



A graphical representation of how to implement AI/ML into multidomain targeting operations using D3A as the framework. (Figure by Maj. Jim Gallagher)

Figure 9. Implementing AI/Machine Learning

enemy course of action and recommend where to concentrate collection assets.

During the detection phase, unmanned systems with object detection models using multiple sensor types will detect enemy equipment and location, which will then be sent to the strike cell. For the delivery phase, a machine learning algorithm built into the Advanced Field Artillery Tactical Data System can recommend the optimal asset to deliver effects to the target. Finally, during the assessment phase of D3A, the LLM can provide measures of performance using object detection to confirm the destruction of enemy equipment and measures of effectiveness to provide analysis of enemy counteraction or reaction based on follow-on intelligence reports. This AI-driven D3A cycle can be highly iterative since it requires minimal time to conduct and minor human involvement, allowing commanders to make better decisions faster to further desynchronize the enemy.

Conclusion

Integrating AI into MDO represents a transformative shift in military strategy and capabilities, offering unprecedented opportunities and challenges. AI's ability to enhance the speed and accuracy of data processing and deliver effects on the battlefield is reshaping the dynamics of modern warfare. However, the adoption of AI systems in MDO is not without its shortfalls. Technical challenges and ethical considerations necessitate a cautious and well-regulated approach to AI integration in the military. The potential for AI systems to be deceived by adversarial tactics, as well as the difficulties in managing and updating AI models across distributed networks, highlights the need for robust, resilient, and lightweight AI solutions tailored to the complexities of the battlefield environment. The evolution of AI capabilities holds the promise of further enhancing the strategic, operational, and tactical advantages in MDO. Yet, these advancements must be accompanied with rigorous testing to address the broader implications of autonomous and semiautonomous weapon systems and LLMs. As we advance into an increasingly AI-integrated future, the focus must remain on developing and deploying AI in a controlled and systematic manner that enhances the U.S. military's capabilities in a multidomain fight while carefully considering and mitigating the associated risks and challenges. ■

Notes

1. István Szabadföldi, "Artificial Intelligence in Military Application—Opportunities and Challenges," *Land Forces Academy Review* 26, no. 2 (1 June 2021): 157–65, https://doi.org/10.2478/raft-2021-0022.

2. Zijie Wang et al., "CNN Explainer: Learning Convolutional Neural Networks with Interactive Visualization," *IEEE Transactions on Visualization and Computer Graphics* 27, no. 2 (February 2021): 1396–406, https://doi.org/10.1109/TVCG.2020.3030418.

3. Zhengxia Zou et al., "Object Detection in 20 Years: A Survey," *Proceedings of the IEEE* 111, no. 3 (March 2023): 257–76, https://doi.org/10.1109/JPROC.2023.3238524.

4. Ibid.

5. Simon Hogue, "Project Maven, Big Data, and Ubiquitous Knowledge: The Impossible Promises and Hidden Politics of Algorithmic Security Vision," in *Automating Crime Prevention, Surveillance, and Military Operations*, ed. Aleš Završnik and Vasja Badalič

(Cham, CH: Springer International Publishing, 2021), 203–21, https://doi.org/10.1007/978-3-030-73276-9_10.

6. Bruce McIntosh, Shashanka Venkataramanan, and Abhijit Mahalanobis, "Target Detection in Cluttered Environments Using Infra-Red Images" (paper presentation, 2020 IEEE International Conference on Image Processing [ICIP], Abu Dhabi, AE, 25–28 October 2020), 2026–30, https://doi.org/10.1109/ICIP40778.2020.9190993.

7. James E. Gallagher and Edward J. Oughton, "Assessing Thermal Imagery Integration into Object Detection Methods on Air-Based Collection Platforms," *Scientific Reports* 13, no. 1 (25 May 2023): Article 8491, https://doi.org/10.1038/s41598-023-34791-8.

8. Ibid.

9. Jianzhong He et al., "BDCN: Bi-Directional Cascade Network for Perceptual Edge Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44, no. 1 (January 2022): 100–13, https://doi.org/10.1109/TPAMI.2020.3007074.

10. James E. Gallagher, Aryav Gogia, and Edward J. Oughton, "A Multispectral Automated Transfer Technique (MATT) for Machine-Driven Image Labeling Utilizing the Segment Anything Model (SAM)," arXiv, 17 February 2024, https://doi.org/10.48550/arXiv.2402.11413.

11. Gallagher and Oughton, "Assessing Thermal Imagery."

12. Mate Krišto, Marina Ivasic-Kos, and Miran Pobar, "Thermal Object Detection in Difficult Weather Conditions Using YOLO," *IEEE Access* 8 (2020): 125459–76, https://doi.org/10.1109/ACCESS.2020.3007481.

13. Gallagher and Oughton, "Assessing Thermal Imagery."

14. Zachary Kallenborn, "Seven (Initial) Drone Warfare Lessons from Ukraine," Modern War Institute, 12 May 2022, https://mwi.westpoint.edu/seven-initial-drone-warfare-lessons-from-ukraine/.

15. Fanlong Zeng et al., "Distributed Training of Large Language Models," *2023 IEEE 29th International Conference on Parallel and Distributed Systems (ICPADS)* (December 2023): 840–47, https://doi.org/10.1109/ICPADS60453.2023.00126.

US ISSN 0026-4148