European Integration Studies, Volume 20, Number 2 (2024), pp. 479-508.
https://doi.org/10.46941/2024.2.18

ZVONKO TRZUN*

Artificial Intelligence and Human-out-of-the-Loop: Is It Time for Autonomous Military Systems?**

ABSTRACT: This paper systematically presents the disruptive technologies that have emerged on the battlefields in recent decades, as well as those that are yet to come. Special attention is given to current technical capabilities: the status of unmanned vehicle development is briefly outlined, focusing primarily on the most prevalent type, unmanned aerial vehicles (UAVs). Additionally, the paper discusses the most common and effective adversarial attack techniques specifically targeting unmanned vehicle technology. The concepts of artificial intelligence (AI), machine learning, deep learning, and convolutional neural networks (CNNs) are introduced. The paper illustrates how CNNs aim to tackle tasks that previously required human intelligence, as well as how the enemy attempts to disrupt the development of CNNs during the crucial training and pattern recognition phase, which is essential for later generalisation. The paper demonstrates the advantages of manned-unmanned teaming as a model that effectively utilises disruptive technologies while simultaneously counteracting the effects of the enemy's measures. Moreover, it analyses the introduction of fully autonomous, AI-driven military systems on the battlefield, outlining the advantages and disadvantages inherent to such a fundamental change. From the evident lack of interest among young people in joining the armed forces to the autonomous systems' potential to save the lives of soldiers and civilians, there are numerous reasons suggesting that this technology could alleviate the burden on human soldiers. However, concerns remain that

* Assist. Prof. Dr.sc., Dr. Franjo Tuđman University of Defense and Security, Croatia. https://orcid.org/0000-0003-3570-9063, zvonko.trzun@sois-ft.hr.
** The research and preparation of this study was supported by the Central European Academy.

autonomous systems may malfunction, potentially reducing rather than increasing the safety of militaries. The paper concludes with recommendations for future steps in the introduction of new technologies, based on their current state of development and the robustness of the AI models they use.

KEYWORDS: Artificial Intelligence, Human-out-of-the-Loop, autonomous military systems, adversary attacks, manned-unmanned teaming.

1. Introduction

Unmanned Military Systems (UMSs) are becoming essential components of military arsenals. Driven by adverse demographic shifts, declining interest in military enlistment, and public aversion to domestic casualties resulting from armed conflicts, UMSs are increasingly deployed on modern battlefields[1]. Once deployed, their outstanding efficiency and the benefits they offer typically justify their substantial initial procurement costs. In essence, unmanned systems are entering and remaining on battlefields around the world, underlining their status as more than just experimental endeavours. Numerous examples of such systems have emerged over the past few decades, to the extent that at present, modern armies cannot be envisioned without them. The groundbreaking moment came with the first actions of the unmanned aerial vehicle (UAV) Predator, recorded at the end of the 20th century. Initially designated as RQ-1 in accordance with the US Air Force's naming conventions, where the letter "Q" is reserved for unmanned aircraft and "R" for reconnaissance missions, this widely utilised UAV underwent a significant transformation in 2002. With the addition of the AGM-114 Hellfire air-to-ground missiles, it was re-designated as MQ-1, signifying its newfound multi-role capabilities.[2] The elegant silhouette, akin to a sailboat, combined with its low weight and large wingspan of 14.8 meters, allowed it to achieve a substantial operational range and endurance in the air for a commendable number of hours.

However, that was just the beginning: the Predator was soon followed by its more powerful successor, the MQ-9 Reaper, boasting a wider operational range and greater endurance (1900 km and 27 h vs. Predator's

[1] Krishnan, 2009, p. 7.
[2] Watts, 2013, p. 18.

1250 km and 24 h), a higher ceiling (50,000 ft. vs. 25,000), increased payload capacity (1750 kg vs. 200 kg) and superior armament (8 AGM-114 Hellfire Missiles, or a combination of Hellfire missiles, GBU-12 Paveway II laser-guided bombs, GBU-38 Joint Direct Attack Munitions, GBU-49 Enhanced Paveway II, or GBU-54 Laser Joint Direct Attack Munitions).[3] The MQ-9 Reaper is also equipped with an enhanced Multi-Spectral Targeting System (MTS), featuring a robust suite of visual sensors for precise targeting. Its MTS-B integrates an infrared sensor, colour and monochrome daylight TV cameras, shortwave infrared camera, laser designator, and laser illuminator.[4] Following the MQ-9 Predator, other sizeable UAVs followed, built on related or independent platforms (for example, the MQ-20 Avenger UAV with jet propulsion and a maximum speed of 720 km/h, the MQ-9B SkyGuardian with an increased 24-meter wingspan and an extended range of 2500 km, or the heavyweight Northrop Grumman's RQ-4 Global Hawk classified as the only HALE USAF unmanned aircraft).

2. Different Classes of Unmanned Vehicles

European industry has also ventured into the development of HALE and MALE UAVs[5], although with less success thus far. However, this does not necessarily imply a negative outcome. Recent military conflicts, particularly the intense battles between Russia and Ukraine, have demonstrated the significant utility of tactical UAVs at a considerably lower cost. In this regard, the European industry has achieved greater success, partly due to reduced technical requirements and development costs, and partly because individual countries were able to develop their own systems instead of engaging in uncertain and often unsuccessful collaborations with other European nations. Notable examples include French Safran's development of the Patroller UAV, Spain's collaboration with Colombia in developing the SIRTAP tactical UAV, and Italy's Leonardo manufacturing the FALCO

[3] US Air Force (2021) MQ-9 Reaper, [Online]. Available at: https://www.af.mil/About-Us/Fact-Sheets/Display/Article/104470/mq-9-reaper/ (Accessed: 18 October 2023).
[4] General Atomics (2020) Lynx Multi-Mode Radar: Surveillance, Tracking, Targeting for Manned and Unmanned Missions, Lynx Datasheet, [Online]. Available at: https://www.ga-asi.com/radars/lynx-multi-mode-radar (Accessed: 5 November 2023).
[5] HALE: High-Altitude Long Endurance UAVs; MALE: Medium-Altitude Long Endurance UAVs.

EVO, among others. These small UAVs feature an endurance exceeding 20 hours, a range of approximately 200 km, a payload capacity of around 200 kg, a ceiling of 6,000 m, and a maximum speed of about 200 km/h. However, there are notable drawbacks, including redundant research and expenditures, resulting in associated costs amounting to a significant 500 million EUR per country. As a result, European attempts to reduce dependence on foreign drone technology through costly capability development projects were unsuccessful.[6] Tactical UAVs, suitable for a wide range of Intelligence, Surveillance, Target Acquisition, and Reconnaissance (ISTAR) missions,[7] are still better solutions for a highly contested airspace, compared to expensive and still vulnerable large MALE and HALE UAVs.

This suggests that it would be beneficial for European nations to undertake joint projects on a more regular basis. However, this is not typical, as each country strives to bolster its technological autonomy, invest in its own manufacturing capabilities, and retain exclusive control over UAV development, including specific requirements. This is particularly evident in the case of the Eurodrone MALE UAV project, which is heavily influenced by conflicting demands from the initiating countries.[8] For instance, Germany prioritised a twin-turboprop configuration for security reasons, whereas France opted for a lighter aircraft. Additionally, Germany has only recently shown interest in developing an armed UAV with attack capabilities, a request initially made by the other founding nations (France, Italy, and Spain).

It would appear that major EU countries are trailing behind in the global market competition, which is dominated by manufacturers from the USA, China, Turkey, and Russia. They have opted to concentrate on creating top-tier MALE UAVs, despite the lessons learned from the Russo-Ukrainian conflict suggesting that quantity often outweighs quality in today's landscape. By committing to developing a single large and costly UAV, which is likely to become outdated by the time of project completion, EU countries continue to fall behind in production and may struggle to set a foothold in the global UAV market. Historical experiences do not favour participation in large-scale EU initiatives either, which tend to result in delayed deliveries of expensive systems with subpar technical capabilities.

[6] Kunertova, 2022, pp. 3-4.
[7] Bartulović, Trzun and Hoić, 2023, p. 87.
[8] Kunertova, 2021, p. 3.

Regrettably, grandiose projects persist while the benefits of miniaturisation, proven effective in recent military conflicts, are disregarded.

As for the unmanned ground, surface, and underwater vehicles (abbreviated UGVs, USVs, and UUVs respectively), their achievements thus far have been modest, primarily due to the challenges of navigating and operating in environments cluttered with obstacles.[9] However, this does not mean that the development of such systems has been halted. For example, the Milrem THeMIS tracked UGV has been built in various iterations, spanning from logistical support vehicles to intelligence, surveillance, and reconnaissance (ISR) variants. Even armed versions of UGVs have been developed capable of carrying various armaments from a 12.7 mm machine gun to a 40 mm grenade launcher. For example, in early 2024, footage was released showing the destructive attack by the Ukrainian Ironclad wheeled UGV on a Russian position. Surface and underwater unmanned systems play an even more significant role in the Russo-Ukrainian war, with the most recent example being on 1 February 2024, when multiple Ukrainian GPS-guided MAGURA V5 USVs attacked and ultimately sank the Russian missile corvette 'Ivanovets' of the Tarantula-III class.

3. Signal Jamming: Obstruction to the Stronger Use of Unmanned Systems

The increasing capabilities of unmanned systems are evident, yet their susceptibility to signal interference cannot be overlooked. Electronic warfare (EW) encompasses the use of a set of powerful tools across three main categories: electronic support (ES), electronic self-protection (EP), and electronic attack (EA). This paper places particular emphasis on EA measures, which involves jamming or other offensive actions aimed at degrading the electromagnetic systems and communications of an adversary. Through EW measures, a less technologically advanced opponent can offset its disadvantage against a more advanced adversary by nullifying the capabilities of its modern systems. As for unmanned vehicles (regardless of their domain), after encountering unbeatable signal interference, they enter idle mode and subsequently circle aimlessly in an attempt to reconnect with the remote pilot.[10]

[9] Zhou et al., 2021, p. 1576.
[10] Smith, 2020, p. 4.

Faced with modern EW measures, unmanned systems actually become more vulnerable the more advanced they are. For example, inertial guidance has its limitations and depends on the accurate operation of accelerometers, so today it is often paired with or even replaced by GPS guidance. However, strong EW defence manages to replace the real GPS signal with a fake one, causing unmanned vehicles to make errors in assessing their location by tens or hundreds of kilometres. Russia was particularly successful in developing its EW capabilities after bad experiences during the 2008 Russo-Georgian War. Based on a sincere analysis and an acknowledgment of the shortcomings of the equipment used at the time, Russia launched its ambitious Armed Forces reform, aiming to have up to 70% new or modernised equipment in the military inventory. This particularly applied to strategic EW systems, which were recognised as an "asymmetric response to the network-centric system of combat operations" on the part of the US and NATO.[11] The Murmansk-BN, a powerful system with a reported range of 5,000 km, capable of the continuous monitoring of electromagnetic activity and intercepting enemy signals with a broad jamming capability, has been recognised as acknowledged the core part of the Russian EW capabilities.

The consequences of such uncompromising modernisation of EW capabilities can be observed today in the Russo-Ukrainian conflict. Regarding the abovementioned GPS spoofing attacks, there are indications that they are deployed across the entire battlefield with significant impact. It is alleged that Russian EW equipment can emit false GPS signals that are an astonishing 500 times stronger than genuine ones.[12]

Generally, the side deploying unmanned vehicles (UVs) seeks to evade the effects of jamming systems and other EW measures by employing more advanced tactics, such as utilising variable frequencies for unmanned vehicle communication with the base station (cognitive radio). Communication between the pilot and the UV is programmed to dynamically change and rapidly select new frequencies to avoid any interruption of data transfer. The algorithm for adjusting transmission parameters continuously analyses the received signal; if adversary interference is detected, changes in transmission parameters, such as

[11] McDermott, 2017, p. 15.
[12] Smith, 2020, p. 4.

frequency, power or modulation are applied. Simultaneously, an alternative frequency range is selected if the current range is deemed unsuitable.[13]

A conflict between two opposing sides where one seeks to disrupt the guidance signal for unmanned vehicles while the other endeavours to evade signal interference poses one of the primary challenges in the wider adoption of UVs. In contemporary conflicts varying in intensity, there is an obvious effort to overcome the defences of the opposing side by deploying a multitude of cheap and disposable robots/drones that attack otherwise well-defended objectives simultaneously. The guidance signal is attempted to be concealed within channels already congested with high data traffic, particularly in urban warfare scenarios. Even highly affordable commercial drones are utilised, with their MAC addresses altered to prevent the identification of the control station (the first six characters of a MAC address denote the manufacturer).[14]

Up to this time, the most prominent instances of employing multiple unmanned vehicles to overload defensive systems took place during the sinking of the cruiser ’Moskva’ and the missile corvette ‘Ivanovets’. In the case of the ’Moskva’, it is alleged that one or two Bayraktar TB-2 UAVs prevented the defensive systems from detecting the incoming ’Neptune’ missile. However, this seems less probable, as the ship's anti-drone and anti-missile defences were provided by two different systems: long-range S-300F (NATO designation: SA-N-6 Grumble) missiles against the Bayraktar and similar slow-moving UAVs, and multi-barrelled AK-630 cannons planned to engage the incoming missiles. On the other hand, during the sinking of the corvette ’Ivanovets’, it appears that six MAGURA V5 USVs easily overwhelmed the relatively weak defence of the Russian ship.

4. Artificial Intelligence and Autonomous Systems

The most effective solution to evade the adversary's EW capabilities could be fully autonomous unmanned vehicles, i.e., vehicles that will advance on the battlefield guided by their own artificial intelligence. The use of autonomous weapon systems offers numerous advantages, ranging from economic and operational to security and humanitarian benefits.[15] From an economic perspective, replacing a destroyed robot or drone is certainly more

[13] Semendiai et al., 2023, p. 731.
[14] Kratky et al., 2020, p. 449.
[15] Monte, 2018, p. 6.

cost-effective than replacing a highly trained, well-equipped soldier. However, this is primarily applicable to Western armies and their warfare strategies, where the adoption of new technologies aims to preserve the lives of their own soldiers (in some other societies, individuals are being seen as easily replaceable assets with minimal economic worth). Autonomous systems can significantly level the playing field between two armies, especially when one possesses a significant numerical advantage in terms of available personnel.

Autonomous systems provide the capability for extremely quick responses to enemy actions. In the event of changes on the battlefield, these systems can swiftly adjust, capitalising on any new opportunities for advancement or promptly reinforcing defences where necessary. The impact of human errors is reduced – a highly significant aspect, especially considering that a significant portion of contemporary accidents, leading to the costly destruction of sensitive equipment, originates from human errors.[16]

From the standpoint of resilience against enemy EW measures, AI-driven systems can continue with combat operations even if the connection with remote pilots is disrupted. In accordance with mission-oriented or mission-type commands, autonomous systems do not require detailed or subsequent instructions once clear objectives are assigned to them. The degree of autonomy depends on the specific system.[17] Semiautonomous systems are often referred to as "human-in-the-loop", where a pilot has to make a positive decision to engage a target. All other actions (such as movement, target tracking, or perimeter monitoring) can be carried out autonomously by such a system. Supervised autonomous systems ("human-on-the-loop") represent the next level, where the robot can autonomously find, identify, and even engage targets, but a pilot monitors the situation and is able to intervene to discontinue the engagement. The highest level of independence is provided by full autonomous weapons, where human pilots are "out-of-the-loop", meaning they have no ability to intervene in the process of weapon engagement. There are also additional classifications of systems based on the level of autonomy achieved.[18] The 'loop' that is mentioned here is actually the OODA loop, which stands for observing,

[16] Wróbel, 2021, p. 9.
[17] Feldman, Dant and Massey, 2019.
[18] Haider, 2021, pp. 14–15.

orienting, deciding, and acting, depending on the current state of the weapon and the target.[19]

Autonomous systems across all three autonomy levels (especially fully autonomous ones) could profoundly alter modern warfare, potentially undermining the current strategies and capabilities of less developed armies. These armies could only reach for the robust EW procedures as a relatively cost-effective asymmetric measure to neutralise the advantages of adversaries with highly sophisticated systems and methods of armed combat.[20]

The enhanced safety of both our troops and civilians is also worth noting. Regarding our forces, it has been previously mentioned that autonomous systems could be deployed in combat operations, either in lieu of soldiers or alongside them, to mitigate the risk of damage. Additionally, in terms of civilian safety, autonomous systems could potentially adhere more strictly to the international humanitarian laws of war, even more reliably than humans, who may be influenced by heightened emotions and stress induced by prolonged fear and uncertainty.[21] Yet, in order for such civilian protection to be effectively realised, it is imperative for autonomous systems to be able to accurately detect civilians and differentiate them from adversary soldiers. Regrettably, AI-driven systems are currently unable to fulfil this task with an adequate level of reliability.

Considering the aforementioned factors, at present it is foreseeable that there will be further advancement in the concept of manned-unmanned teaming (MUM-T). This concept emphasises a team encompassing multiple units, with the human-operated unit retaining a central role, while additional AI-driven units serve for support and protection. Over time, these AI-driven units are likely to be granted increasing levels of autonomy and to be tasked with more complex assignments. However, the central unit should always remain under human control, ensuring oversight over the entire team. The MUM-T approach is moving towards a model of ’human-on-the-loop’ supervised autonomy, wherein AI-driven units can autonomously perform a significant part of their tasks, thereby relieving humans from routine supervisory duties such as movement or obstacle avoidance. Nonetheless, human intervention remains crucial for decisions about whether certain actions should proceed or be halted.

---

[19] Morgan et al., 2020, p. 12.
[20] McDermott, 2017, p. 3.
[21] Monte, 2018, p. 162.

The emergence of the MUM-T concept is expected to remain a prominent trend for the foreseeable future, spanning over years or even decades. AI models will require thorough testing and refinement, raising questions about the feasibility of granting them full autonomy ('human-out-of-the-loop'), given the potential for numerous incidents and collateral damage. In parallel with technological progress, there must be a concerted effort to develop a suitable legal framework, which may involve amendments to international humanitarian law.

As for the "human-out-of-the-loop" (HOOTL) concept, it indeed offers a number of advantages. It provides unprecedented efficiency and speed, scalability (these systems can handle large-scale operations without the limitations of human attention span and fatigue), increased safety for the implementing side, and significant cost reduction. HOOTL fully autonomous weapons and surveillance systems could operate independently in complex and potentially hostile environments. Nevertheless, there are also many challenges and risks associated with such systems, where safety and reliability are paramount. Errors or malfunctions can lead to catastrophic consequences. As AI and machine learning technologies advance, the potential for HOOTL systems to become more prevalent increases.

5. Techniques and Tools of Artificial Intelligence

Artificial Intelligence (AI) represents a specific field of computer science that deals with creating systems capable of performing tasks that typically require human intelligence. One of the key tools in the field of AI is machine learning (ML), which enables computers to learn from experience without explicit programming. ML is based on the concept of algorithms that analyse data, identify patterns in those data, and use those patterns to make decisions or predictions. Examples of ML applications span from image and speech recognition to product recommendations and data analysis.[22]

Machine learning is characterised by its capability to automatically enhance system performance through experience. Instead of manual rule definition by programmers, ML algorithms utilise data to discern implicit patterns and regularities, applying acquired knowledge to novel,

[22] Wang and Siau, 2019, pp. 61-63.

unencountered situations.[23] There are three primary types of ML: supervised learning, unsupervised learning, and reinforcement learning. In supervised learning, algorithms are trained on labelled data with correct answers, with the aim to generalise learned patterns to new, unlabelled data. Unsupervised learning involves analysing data that lack labelled correct answers, with algorithms tasked with discovering hidden patterns and structures, such as clustering similar items or reducing dimensionality. In reinforcement learning, algorithms interact with the environment, adjusting their strategies based on feedback to maximise rewards or minimise penalties.[24]

A special and widely applicable subtype of ML is deep learning (DL), used for its ability to learn from highly complex datasets. Deep learning also uncovers patterns in data but employs different techniques. Both methodologies (DL and ML) start out with training using sample data and models, during which they establish relevant connections between different data points. Following this, they undergo an optimisation process to ascertain the most precise weighted values among these connections and to ensure that the model aligns as closely as possible with the data.

DL employs artificial neural networks with numerous layers, hence the term "deep". These networks can recognise intricate patterns within data, allowing them to address highly complex tasks.[25] Rather than manually defining features or rules, deep neural networks learn implicit patterns and structures through layers of data transformations. Each layer processes input data and generates output features, which then serve as input for subsequent layers, allowing for a progressive abstraction and broader generalisation of the data.[26]

Some of the most commonly utilised DL networks are convolutional neural networks (CNNs) and recurrent neural networks (RNNs). CNNs are particularly efficient in analysing images and video content by utilising convolutional layers to extract local features and reducing the dimensionality of input data. Conversely, RNNs are adept at handling sequential data such as text or time series, utilising recurrent connections between neurons to model temporal dependencies.[27]

---

[23] Janiesch, Zschech and Heinrich, 2021, p. 686.
[24] Carleo et al., 2019, p. 045002-5.
[25] Bengio, Lecun and Hinton, 2021, p. 60.
[26] Mu and Zeng, 2019, p. 1745.
[27] Janiesch, Zschech and Heinrich, 2021, pp. 688-690.

CNNs have achieved remarkable results in areas such as object recognition, image classification, face detection, medical diagnostics, and other domains where visual data analysis is utilised. The main characteristic of CNNs is the use of convolutional layers, alongside which CNNs typically involve pooling layers that serve to reduce the dimensionality and computational complexity of the model. The aim is to aggregate and summarise information from convolutional layers, thereby facilitating the further processing and interpretation of features.

A key advantage of CNNs is their ability to automatically learn hierarchical features from input data. In this sense, CNNs seek to simulate the functioning of the central nervous system of living organisms, namely the brain. Similar to our biological nervous system, CNNs consist of simple processing units whose task is mutual communication through a high number of connections.[28] Instead of manually defining features or patterns, CNNs use data-driven learning through an iterative process of optimising network weights to minimise prediction errors. An activation function (often referred to as a transfer function) is then used for further information transfer. Some of the most common ones are the threshold function, the piecewise linear function, and the sigmoid function.

The technique of using CNNs has attained outstanding results in many tasks, at times surpassing human capabilities. It is applied in image recognition, object detection and segmentation, medical diagnostics, natural language translation, time series analysis, and much more – including autonomous driving through the analysis of geospatial data. However, it is important to emphasise that the level of accuracy and reliability of these techniques still varies depending on the data presented and the quality of the training process.

6. Problems and Limitations of AI Training

Below are some of the most common issues encountered with the techniques discussed above.

6.1. Data Bias

If the training dataset is not sufficiently diverse or representative, the algorithm may learn biased patterns and become unbalanced in its predictions. Data bias, also known as dataset imbalance, is caused by a

[28] Li et al., 2021, pp. 6999-7002.

situation where certain classes or categories have a greater number of examples in the dataset compared to other classes. This phenomenon often occurs in real-world datasets due to natural variations or irregularities in the data collection process and can result in unfair models that prefer dominant classes, while neglecting or misclassifying less represented ones.[29] For example, in a dataset aimed at recognising armoured vehicles, there might be more images of Abrams tanks than images of other tanks. If a CNN is trained on such a dataset, there is a risk that the model will recognise Abrams tanks better than other armoured vehicles, which will subsequently be recognises with significantly lower reliability.

Employing biased algorithms in autonomous weapons systems would negatively impact already marginalised groups. The solution to this problem involves collecting a larger and more diverse dataset, along with additional data collection for less represented classes, and applying techniques such as data augmentation (generating new examples from existing data) or adjusting weights in learning algorithms to account for class imbalances in the sample.

6.2. Overfitting

Overfitting is a common problem in the context of CNNs. When a CNN becomes too tailored to the training dataset, it can lose the ability to generalise to new data. Solutions for overfitting include using regularisation techniques such as dropout, early stopping, and gradient normalisation.

Overfitting can occur for various reasons. One of the main causes is the complexity of the model. If the model is too complex and has too many parameters relative to the amount of available data, it may learn overly complex patterns that are not necessarily relevant to the general population of data. The model develops excessive adaptations to the sample used during training, while losing the ability to generalise to new data acquired later. Unlike human problem solving, which is inherently flexible and capable of adapting to new and diverse challenges, machine-learning systems are usually not transferable to entirely different problem contexts.[30]

Overfitted models have poor generalisation ability with regard to new data, resulting in poor performance in real-world applications. For example, if an overfitted model is used for image classification, incorrect predictions

[29] Ntoutsi et al., 2020, pp. 4-5.
[30] Surden, 2021, p. 175.

may ensue when the model is applied to images that were not present in the training dataset.

Fortunately, there are various strategies for addressing overfitting (if a larger training dataset is not available). One of the most common strategies is regularisation, which involves adding additional constraints to the model to prevent overfitting. This can include techniques such as dropout, where certain neurons are randomly excluded during training, as well as cross-validation and early stopping.

6.3. Scarcity of Data

Under particular circumstances, acquiring the sufficient volume of data for training a CNN may present challenges, particularly in cases involving constrained datasets, such as those pertinent to medical diagnostics. The process of data collection for scientific inquiry can similarly involve significant costs, time investments, or ethical considerations. Moreover, impediments of a technical or legal nature might obstruct access to extant datasets. Irrespective of the underlying factors contributing to these obstacles, the scarcity of data can curtail the CNN's capability to learn general patterns and structures.[31]

An effective strategy to address this scenario involves employing transfer learning methodologies, wherein the model undergoes training on a comparable yet more expansive dataset. Along with regulating the quantity, it is imperative to oversee the quality, particularly the representativeness, of the training data. This entails processes such as filtering, cleansing, and normalising the data to eliminate any problematic or incongruous instances.[32]

6.4. Interpretability

The interpretability of CNNs, or the ability to understand and explain their predictions, also poses a significant challenge. Given that CNNs are complex models with numerous parameters, it is difficult to discern the features or patterns utilised by the model to make decisions.

Interpretability could prove to be a crucial aspect in the context of public trust in AI, as it helps understand why models have made certain decisions and how they have arrived at their predictions. Public trust is particularly vital in critical domains such as medical diagnostics, finance,

[31] Janssen et al., 2020, p. 2.
[32] Bansal, Sharma and Kathuria, 2022, p. 8.

and legislative issues[33] — but perhaps most notably in the realm of military decision-making.

There are several approaches to interpretability in machine learning. One of them is feature visualisation, where techniques like heatmaps and saliency maps are employed to display the relevant features of input data that have influenced the model's final decision. Additionally, attribution methods such as LIME (Local Interpretable Model-agnostic Explanations) approximate any black-box ML model to a local, interpretable model.

6.1.1. Introducing Noise into Data During Wartime

In wartime conditions, the adversary will likely undertake all available actions to disrupt the process of training AI models or to corrupt established connections. Introducing noise into field data poses a significant problem in the realm of data analysis and machine learning. Noise can be defined as unwanted additions to or disturbances in data and can be introduced from various sources. Noise can compromise the accuracy and reliability of data analysis or model predictions. For example, in image recognition or object detection in images, the presence of noise can lead to incorrect classifications or inaccurate predictions. Furthermore, noise can reduce the interpretability of analysis results by making it difficult to distinguish relevant signals from unwanted interference.[34] Finally, noise can increase the complexity of the model and consume resources for data processing and learning.

Resolving the problem of noise in input data requires the application of various strategies and techniques. One possible approach is the application of data filtering and cleaning, where algorithms are used to detect and remove noise. This technique may involve the use of different filters such as median filters or averaging. Another possibility is to apply techniques which reduce the model's sensitivity to noise. This might include employing robust algorithms that are more resilient to data noise or utilising regularisation techniques to avoid overfitting on data corrupted by noise.[35]

The strategy of introducing noise can be viewed as a type of electronic warfare or, alternatively, a form of tactical deception. Injecting noise into the visual identification of equipment by adversarial systems can be

[33] Rodrigues, 2020, p. 2.
[34] Xiong et al., 2006., pp. 305-307.
[35] Gupta and Gupta, 2019, pp. 471-472.

executed through diverse methods, contingent upon the precise technical attributes of the image recognition system being used. Here are a few possible scenarios.

The first is image manipulation, which means altering or distorting the appearance of one's equipment to make it less recognisable to adversary systems. This may include adding false details, altering colours or textures, or even completely changing the visual shape to deceive image recognition algorithms. The problem of image manipulation is particularly significant with the advent of deep-fake technology. Innovative tools are being developed to detect such manipulations and uncover genuine information,[36] but at the same time, new methods for even more sophisticated image and video manipulation are being constantly revealed.[37]

The second is masking, which means employing camouflage techniques to hide equipment. This can entail using colours and patterns that blend seamlessly with the surroundings. Additionally, natural cover or artificially created shapes may be utilised to integrate the equipment into the environment, making it less conspicuous to sensors. Furthermore, equipment can be coated with reflective materials to disrupt enemy IC or laser sensors.

The third is distorting sensor data, which implies disrupting the operation of sensors or cameras using flashes, laser devices, or other devices that could interfere with or overload enemy sensors.

And finally, there is the injection of false data, namely introducing distorted facts or images into the training set of the opponent's system, leading it to draw incorrect conclusions. This can be done by sending false signals or data through electronic communication channels, or even by hacking the opponent's system while it is still in the training phase.[38,39]

Employing such tactics carries significant implications, including potential ethical and legal ramifications. While the methods described may indeed disrupt the adversary's recognition systems, they also pose risks of unintended consequences or misinterpretations of the battlefield situation. Specifically, such tactics could lead to false negatives, where AI systems fail to identify the adversary's assets or combatants accurately, but also lead to false positives, where the adversary's AI wrongly identifies civilians and

[36] Lee et al., 2023, pp. 3-4.
[37] Zhang, Li and Chang, 2024, p. 4.
[38] Tufail, Batool and Sarwat, 2021, p. 3.
[39] Gong and Wang, 2023.

their vehicles or structures as military targets. Given these risks, the legitimacy of such tactics is subject to scrutiny. The use of noise can be interpreted as a form of unfair combat or a violation of international rules of warfare, especially if it results in unjustified civilian casualties or unnecessary destruction.

In their paper "Intriguing Properties of Neural Networks", Szegedy et al. (2014)[40] introduced the concept of an "adversarial example", which refers to an example created with the aim of manipulating or inducing errors in deep learning models. The authors acknowledge that deep neural networks (DNNs) are 'powerful learning models that achieve excellent performance on visual and speech recognition problems,' but they also point out two counter-intuitive properties of deep neural networks. The first is a significant question regarding the conjecture that neural networks disentangle variation factors across coordinates. The second is related to the stability of neural networks with respect to small perturbations to their inputs. Unlike intuitive thinking, DNNs (which otherwise generalise well on the task of object recognition) may react even to very small perturbations, carefully crafted so that the DNN completely misidentifies the object category in the presented image.

Adversarial attacks involve making slight alterations to input data, introducing changes so subtle that they are practically imperceptible to the human eye. On the other hand, DNNs can become "confused" and produce erroneous object detections on images manipulated by adversaries. An illustrative instance of such an attack is the image of a panda, initially identified by a DNN with a confidence of 57.7%. However, after injecting noise into the image, the DNN incorrectly classified the object as a gibbon with an exceedingly high confidence level of 99.3%.[41] Similarly, 3D-printed toy turtles were persistently misidentified as rifles by the targeted AI.[42]

Some methods can enhance the resilience of DNNs against attacks, like expanding capacity (by incorporating more connections into a DNN) and adversarial training (training DNNs where each input is adjusted by a synthetic adversary before being processed by the network). While these approaches enable DNNs to maintain some level of accuracy in the face of attacks, they are particularly resource-intensive, demanding substantially

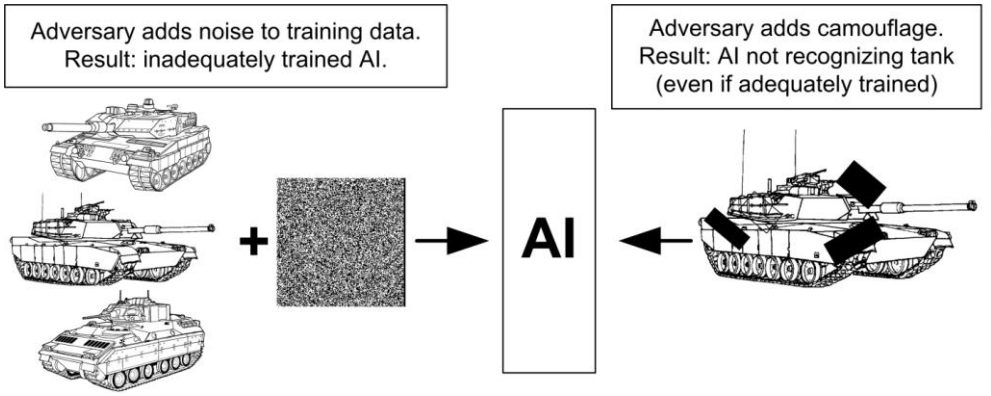
[40] Szegedy et al., 2014, p. 2.
[41] Goodfellow, Shlens and Szegedy, 2014, p. 3.
[42] Athalye et al., 2018, p. 284.

more storage and computational resources. Consequently, they become highly impractical for everyday usage.[43]

The adversary attacks described above involve modifications to images that are almost imperceptible to the human eye before being presented to the DNN. However, if more significant alterations are applied (physical adversarial perturbations), the outcomes become even more striking. This relates to the previously discussed ability to deceive image recognition models, as numerous studies have shown the effects of alterations that the human brain notices but is not deceived by them, while AI algorithms struggle to interpret them accurately. These studies demonstrate that even by adding small stickers to the surface of an object that the attacker seeks to conceal (e.g., a military vehicle) a significant number of misidentifications ensue.[44] Equally vulnerable are today's commercially available autonomous driving models that can be easily disguised as changes to traffic signs and fake obstacles by malicious attackers.[45] The placement of counterfeit lane markers is particularly dangerous, as it can easily cause vehicles to veer off their intended path of travel. Adversary attacks have been presented in Figure 1.

[43] Gilles, 2020, p. 19.
[44] Brown et al., 2017, pp. 4-5.
[45] Eykholt et al., 2018, p. 1626.

Figure 1 *Different kinds of adversary attacks: adding noise to training data (left) and using camouflage (right).*



Source: original author's work.

Different strategies are being taken into consideration in the literature as responses to adversarial attacks. These countermeasures can be broadly classified into three main categories: 1) gradient masking, which aims to conceal or obscure the gradient information of the classifier, 2) robust optimisation, which involves the re-learning of the parameters of a DNN classifier, and 3) adversarial examples detection, which focuses on identifying adversarial examples and preventing them from being fed into the classifier.[46] However, considering that the attacker always plays an active role, meaning they are the first to discover the new methods of provoking false detections to which the defending side must then find a response, we can conclude that the advantage lies on their side.

7. The Analysis: Perceived Benefits and Drawbacks of AI-driven Military Systems

Although autonomous, AI-driven military systems such as UAVs, UGVs, unmanned surface vehicles (USVs), and unmanned underwater vehicles (UUVs) have yet to see extensive implementation on battlefields, the potential they hold motivates military leadership to continually push for their accelerated development. This chapter will present a brief analysis,

[46] Xu et al., 2020, p. 161.

focusing on the major benefits and drawbacks of fully autonomous military systems as they are perceived today.

7.1. Advantages

Reduced Risk to Human Lives: One of the most compelling arguments in favour of autonomous military systems is their ability to minimise the risk to human militaries. By deploying unmanned vehicles and drones, combatants can conduct reconnaissance, surveillance, and even combat operations without endangering soldiers' lives. Autonomous systems can navigate through unsafe terrain, detect and disarm explosives, and engage enemy targets. Autonomous systems can also be utilised for logistical support and supply delivery, further moderating the exposure of human personnel to potential threats.

Enhanced Situational Awareness: Autonomous systems equipped with advanced sensors and surveillance capabilities provide real-time situational awareness to military commanders and personnel. This comprehensive understanding of the battlefield facilitates s strategic decision-making while minimising the need for soldiers to physically scout enemy positions or gather intelligence in dangerous areas.

Reduced Psychological Impact: Warfare can have significant psychological effects on soldiers, including post-traumatic stress disorder (PTSD) and other mental health issues. By leveraging autonomous systems for combat and support operations, military forces can potentially reduce the psychological burden on human personnel, sparing them from the trauma associated with direct engagement in conflict.

Humanitarian Considerations: By employing autonomous systems to carry out missions with precision and efficiency, militaries can strive to minimise civilian casualties and collateral damage, thereby upholding the principles of proportionality and distinction in an armed conflict. Autonomous systems can execute missions with minimal deviation from objectives, and their precision is particularly valuable in targeted strikes against high-value targets surrounded by civilians.

The ability to operate 24/7: Unlike human soldiers who require rest and sleep, autonomous systems can operate continuously, with the capacity of providing persistent surveillance and monitoring. This enables militaries to maintain constant vigilance over large areas for extended periods, improving situational awareness and response times.

Cost-Efficiency: While the initial development and procurement costs of autonomous systems can be high, they often prove cost-effective in the long run. Compared to retaining large standing armies or deploying manned aircraft, autonomous systems are more affordable to deploy and maintain, particularly in prolonged conflicts.

7.2. Disadvantages

The Risk of the Autonomous System Executing Incorrect Actions: Perhaps the most significant drawback of autonomous military systems is the risk that they may not function as intended. Despite their precision, autonomous systems are not immune to errors or malfunctions. Software glitches, communication failures, or misinterpretation of data can lead to unintentional consequences, including civilian casualties or friendly fire incidents. The potential for these systems to malfunction raises significant concerns regarding their reliability and safety. In the preceding sections, we have outlined the methods through which our adversaries might intervene and disrupt the AI training process. In such an event, autonomous systems, although designed to target specific objectives with precision, may mistakenly identify and engage non-combatants or civilian infrastructure. Miscommunication, faulty identification algorithms, or inaccurate situational awareness may also lead to friendly fire incidents.

Technical Failures: Just as in case of any other technical object, autonomous systems are also susceptible to technical failures, including hardware malfunctions, software glitches, and sensor errors. These failures may be caused by manufacturing defects, environmental factors, or wear and tear over time, leading to disruptions in operation and potential mission failure. Accessing and servicing autonomous systems deployed in remote or hostile environments can pose logistical challenges, potentially leading to delays in maintenance and reduced system availability. On the other hand, poor reliability erodes trust and confidence in autonomous systems among operators, commanders, and stakeholders. Concerns about the system's ability to perform reliably under operational conditions may lead to hesitancy in relying on autonomous capabilities, resulting in a reluctance to fully integrate these systems into military operations.

Ethical and Moral Concerns: Concerns about accountability, decision-making ethics, and the potential for autonomous weapons to violate international humanitarian law raise profound moral questions. The lack of

human oversight in critical decision-making processes can lead to unintentional consequences and ethical breaches.

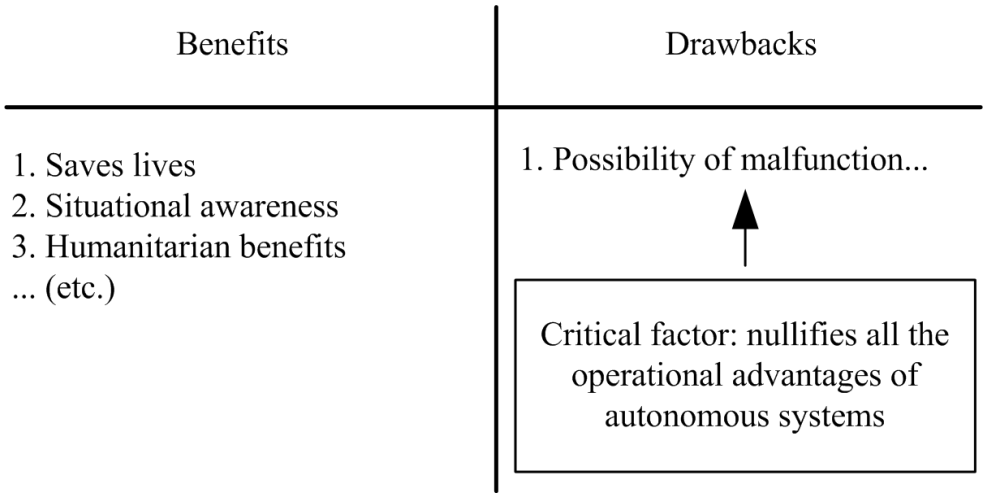
Lack of Emotional Intelligence and Contextual Understanding: Autonomous systems lack the emotional intelligence and contextual understanding of human soldiers. They may struggle to interpret complex social and cultural dynamics, leading to misjudgements or inappropriate responses in sensitive situations. Additionally, the absence of human intuition and empathy can hinder their ability to make nuanced decisions in dynamic and unpredictable environments.

Legal and Regulatory Challenges: The existing legal frameworks governing the use of autonomous vehicles and weapons are insufficient to address the complex challenges autonomous systems pose. Questions regarding accountability, liability, and compliance with international humanitarian law remain unresolved. Establishing clear regulations and norms for the use of autonomous military systems is essential to mitigate the risks associated with their deployment.

7.3. Conclusion of the Analysis

The preceding comparison clearly illustrates the numerous benefits of autonomous systems, with the greatest one certainly being the potential to save the lives of soldiers and civilians in war zones. Considering these arguments, the implementation of AI-driven systems, as swiftly and extensively as possible, enjoys almost unquestionable support. However, what alters the conclusions of the analysis is the risk that autonomous systems may fail to fulfil their mission or even commit errors so severe that they could endanger friendly troops and civilians (Figure 2).

Figure 2 *The malfunctioning of a Figure 1: Different kinds of adversary attacks: adding noise to training data (left) and using camouflage (right). As a result of this threat, we believe that the deployment of fully autonomous systems is still premature, at least until issues stemming from sensor errors, enemy electronic warfare, and insufficiently robust AI models are addressed.*



Source: original author's work.

Several authors maintain that autonomous weapons need to be used along with intelligible human control to comply with legal and ethical norms – in other words, the use of weapons without meaningful human control should be prohibited. Fully autonomous weapons systems do not allow a human to make a legal and moral judgment as to whether the effects of an attack are acceptable. A treaty that would restrict the use of autonomous military systems should not be built around specific existing technologies but rather based on the idea of how technology may evolve and how it could be used in the future.

Controlling lethal autonomous weapons systems (LAWS) is imperative to ensure adherence to international law, particularly the principles of distinction, proportionality, and precautions in attacks as delineated by International Humanitarian Law (IHL). Human judgment plays a pivotal role (and should not be excluded from the decision-making chain) in ensuring that the potential deployment of LAWS is consistent with

international legal norms and IHL standards. Consequently, there is a critical need for maintaining and enhancing human-machine interaction, where human decision-making continues to hold superiority over decisions made exclusively by AI.

8. Summary

In this paper, we have presented an assessment of the development of disruptive technologies. These technologies vary in terms of their capabilities, acceptance, and dissemination. Using unmanned aerial vehicles (UAVs) and unmanned ground vehicles (UGVs) as examples, the assessment reveals that UAVs have already established themselves as widely adopted, high-capability technology, while UGVs are still searching for their place within global armed forces.

Moreover, the acceptance and applicability of AI within military systems have been thoroughly assessed. Despite considerable progress in the processing of images and real-time video transmissions, the potential for the misidentification of observed objects remains significant. Granting complete autonomy to current AI-driven systems could also entail a notable risk of inadvertent engagement with civilian or neutral targets. Such occurrences may arise from the absence of adequate sensors or models, or as a result of adversary attacks.

The implications of the aforementioned errors in AI-powered military systems diverge in severity, though none of these can be dismissed as insignificant. For example, AI may incorrectly classify an enemy vehicle or weapon, leading to the selection of inappropriate weaponry or tactical manoeuvres. Additionally, mistaking a friendly vehicle for an adversary could result in incidents of friendly fire and fratricide. In view of the attained capabilities and vulnerability to adversary attacks, it currently appears unfeasible for AI to effectively monitor the movements of multiple entities and swift changes on the battlefield while maintaining the requisite high level of situational awareness. If granted complete autonomy, AI-driven military systems would need to accurately and flawlessly distinguish between friendly troops, enemy combatants, and unarmed civilians. AI should be able to discern whether a person is carrying a weapon or any other item and adjust its responses accordingly, with only the highest level of reliability deemed acceptable.

Considering the stated factors, further advancement in manned-unmanned teaming (MUM-T) is predictable. This concept involves a team where the human-operated unit plays a pivotal role, with AI-driven units providing support and protection. These AI units will likely gain more autonomy and take on complex tasks over time, but the central unit will always remain under human control. The MUM-T approach is evolving towards "human-on-the-loop" supervised autonomy, where AI units manage routine tasks while humans make critical decisions. This trend is expected to persist for years or decades, with thorough testing and a legal framework required to ensure safety and compliance with international humanitarian law.

In the final chapter, an analysis was conducted regarding the advantages and disadvantages of fully autonomous systems. Arguments supporting the potential to save the lives of soldiers and civilians serve as the primary motivation for the eventual deployment of such technical units, ideally in significant numbers. However, numerous still-unresolved issues, ranging from hardware and software imperfections to insufficient resilience against enemy attacks, warrant caution. With the declining number of young people willing to enlist in the military, it is almost certain that autonomous systems will eventually assume a significant share of tasks currently reliant on human soldiers. Nevertheless, insistence on such a fundamental transition must be tempered until the aforementioned issues have been addressed, as unsuccessful experiments will be paid for in blood and human lives.

Bibliography

[1] Athalye, A., Engstrom, L., Ilyas, A., Kwok, K. (2018) 'Synthesizing Robust Adversarial Examples', in Dy, J., Krause, A. (eds) *Proceedings of the 35th International Conference on Machine Learning. PMLR (Proceedings of Machine Learning Research),* pp. 284–293. [Online]. Available at: https://proceedings.mlr.press/v80/athalye18b.html (Accessed: 18 October 2023).

[2] Bansal, M. A., Sharma, D. R. Kathuria, D. M. (2022) 'A systematic review on data scarcity problem in deep learning: solution and applications', *ACM Computing Surveys (CSUR),* 54(10s), pp. 1–29; https://doi.org/10.1145/3502287.

[3] Bartulović, V., Trzun, Z., Hoić, M. (2023) 'Use of Unmanned Aerial Vehicles in Support of Artillery Operations', *Strategos,* 7(1), pp. 71–92.

[4] Bengio, Y., Lecun, Y. Hinton, G. (2021) 'Deep learning for AI', *Communications of the ACM,* 64(7), pp. 58–65; https://doi.org/10.1145/3448250.

[5] Brown, T., Mané, D., Roy, A., Abadi, M., Gilmer, J. (2017) '*Adversarial Patch'.* arXiv:1712.09665; https://doi.org/10.48550/arXiv.1712.09665.

[6] Carleo, G., Cirac, I., Cranmer, K., Daudet, L., Schuld, M., Tishby, N., Vogt-Maranto, L., Zdeborová, L. (2019) 'Machine learning and the physical sciences', *Reviews of Modern Physics,* 91(4), pp. (045002)1-39; https://doi.org/10.1103/RevModPhys.91.045002.

[7] Eykholt, K., Evtimov, I., Fernandes, E., Li, B., Rahmati, A., Xiao, C. (2018) 'Robust Physical-World Attacks on Deep Learning Visual Classification', in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1625–1634; https://doi.org/10.1109/CVPR.2018.00175.

[8] Feldman, P., Dant, A. Massey, A. (2019) *'Integrating artificial intelligence into weapon systems'*, arXiv preprint arXiv:1905.03899 [Preprint].

[9] Gilles, J. (2020) *The lottery ticket hypothesis in an adversarial setting*. Massachusetts: Massachusetts Institute of Technology.

[10] Gong, Z., Wang, W. (2023) 'Adversarial and clean data are not twins', *Proceedings of the Sixth International Workshop on Exploiting Artificial Intelligence Techniques for Data Management,* pp. 1–5; https://doi.org/10.1145/3593078.3593935.

[11] Goodfellow, I. J., Shlens, J., Szegedy, C. (2014) 'Explaining and Harnessing Adversarial Examples', *CoRR,* abs/1412.6; https://doi.org/10.48550/arXiv.1412.6572.

[12] Gupta, S., Gupta, A. (2019) 'Dealing with noise problem in machine learning data-sets: A systematic review', *Procedia Computer Science,* 161, pp. 466–474; https://doi.org/10.1016/j.procs.2019.11.146.

[13] Haider, A. (2021) 'Introduction', in Willis, M., Haider, A. (eds) *A Comprehensive Approach to Countering Unmanned Aircraft Systems.* Kalkar, Germany: Joint Air Power Competence Centre, pp. 14–15; https://doi.org/10.1007/978-3-030-67341-3_1.

[14] Janiesch, C., Zschech, P.,Heinrich, K. (2021) 'Machine learning and deep learning', *Electronic Markets,* 31(3), pp. 685–695; https://doi.org/10.1007/s12525-021-00475-2.

[15] Janssen, M., Brous, P., Estevez, E., Barbosa, L. E., Janowski, T. (2020) 'Data governance: Organizing data for trustworthy Artificial Intelligence', *Government information quarterly,* 37(3), 101493; https://doi.org/10.1016/j.giq.2020.101493.

[16] Kratky, M., Minarik, V., Sustr, M., Ivan, J. (2020) 'Electronic Warfare Methods Combatting UAVs', *Advances in Science, Technology and Engineering Systems Journal,* 5(6), pp. 447–454; https://doi.org/10.25046/aj050653.

[17] Krishnan, A. (2009) *Killer Robots: Legality and Ethicality of Autonomous Weapons.* Surrey, UK: Ashgate Publishing Limited.

[18] Kunertova, D. (2021) 'European Drone Clubs Stall Strategic Autonomy', *CSS Policy Perspectives*, 9(5), pp. 1-4.

[19] Kunertova, D. (2022) 'The Ukraine Drone Effect on European Militaries', *CSS Policy Perspectives,* 10(15), pp. 1-4; https://doi.org/10.3929/ethz-b-000584078.

[20] Lee, J., Jeon, S., Park, Y., Chung, J., Jeong, D. (2023) '*A Forensic Methodology for Detecting Image Manipulations'*, arXiv preprint arXiv:2308.04723 [Preprint].

[21] Li, Z., Liu, F., Yang, W., Peng, P., Zhou, J. (2021) 'A survey of convolutional neural networks: analysis, applications, and prospects', *IEEE transactions on neural networks and learning systems,* 33(12), pp. 6999–7019; https://doi.org/10.1109/TNNLS.2021.3084827.

[22] McDermott, R. (2017) *Russia's Electronic Warfare Capabilities to 2025: Challenging NATO in the Electromagnetic Spectrum*. Talinn: International Centre for Defence and Security.

[23] Monte, L. Del (2018) *Genius Weapons.* New York: Prometheus Books.

[24] Morgan, F. E., Boudreaux, B., Lohn, A. J., Ashby, M., Curriden, C., Klima, K., Grossman, D. (2020) *Military Applications of Artificial Intelligence: Ethical Concerns in an Uncertain World.* Santa Monica, CA: RAND Corporation, https://doi.org/10.7249/RR3139-1.

[25] Mu, R., Zeng, X. (2019) 'A review of deep learning research', *KSII Transactions on Internet and Information Systems (TIIS),* 13(4), pp. 1738–1764; https://doi.org/10.3837/tiis.2019.04.001.

[26] Ntoutsi, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdl, W., Vidal, M.-E., Ruggieri, S., Turini, F., Papadopoulos, S., Krasanakis, E., Kompatsiaris, I., Kinder-Kurlanda, K., Wagner, C., Karimi, F., Fernandez, M., Alani, H., Berendt, B., Kruegel, T., Heinze, C., Broelemann, K., Kasneci, G., Tiropanis, T., Staab, S. (2020) 'Bias in data-driven artificial intelligence systems—An introductory survey', *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(3), p. e1356; https://doi.org/10.1002/widm.1356.

[27] Rodrigues, R. (2020) 'Legal and human rights issues of AI: Gaps, challenges and vulnerabilities', *Journal of Responsible Technology*, 4, p. 100005; https://doi.org/10.1016/j.jrt.2020.100005.

[28] Semendiai, S., Tkach, Y., Shelest, M., Korchenko, O., Ziubina, R., Veselska, O. (2023) 'Improving the Efficiency of UAV Communication Channels in the Context of Electronic Warfare', *International Journal of Electronics and Telecommunications*, 69(4), pp. 727–732; https://doi.org/10.24425/ijet.2023.147694.

[29] Smith, P. (2020) *Russian Electronic Warfare: A Growing Threat to U.S. Battlefield Supremacy.* American Security Project.

[30] Surden, H. (2021) 'Machine learning and law: An overview', *Research Handbook on Big Data Law*, pp. 171–184; https://doi.org/10.4337/9781788972826.00014.

[31] Szegedy, C., Zaremba, W., Sutskever, I., Bruna, J., Erhan, D., Goodfellow, I., Fergus, R. (2014) *'Intriguing properties of neural networks'*, ArXiv [Preprint].

[32] Tufail, S., Batool, S., Sarwat, A.I. (2021) 'False data injection impact analysis in AI-based smart grid', *SoutheastCon 2021*. pp. 1–7; https://doi.org/10.1109/SoutheastCon45413.2021.9401940.

[33] Wang, W., Siau, K. (2019) ‘Artificial intelligence, machine learning, automation, robotics, future of work and future of humanity: A review and research agenda’, *Journal of Database Management (JDM),* 30(1), pp. 61–79; https://doi.org/10.4018/JDM.2019010104.

[34] Watts, B. (2013) *The Evolution Of Precision Strike.* Washington DC: Center for Strategic and Budgetary Assessments.

[35] Wróbel, K. (2021) ‘Searching for the origins of the myth: 80% human error impact on maritime safety’, *Reliability Engineering & System Safety,* 216; https://doi.org/10.1016/j.ress.2021.107942.

[36] Xiong, H., Pandey, G., Steinbach, M., Kumar, V. (2006) ‘Enhancing data analysis with noise removal’, *IEEE transactions on knowledge and data engineering,* 18(3), pp. 304–319; https://doi.org/10.1109/TKDE.2006.46.

[37] Xu, H., Ma, Y., Liu, H.-C., Deb, D., Liu, H., Tang, J.-L., Jain, A. K. (2020) ‘Adversarial attacks and defenses in images, graphs and text: A review’, *International journal of automation and computing*, 17, pp. 151–178.

[38] Zhang, Z., Li, M., Chang, M.-C. (2024) ‘A New Benchmark and Model for Challenging Image Manipulation Detection’, *Proceedings of the AAAI Conference on Artificial Intelligence,* 38(7), pp. 7405-7413; https://doi.org/10.1609/aaai.v38i7.28571.

[39] Zhou, X., Xiang, Y., Youmin, Z., Yangyang, L., Xiaoyan, P. (2021) ‘Trajectory Planning and Tracking Strategy Applied to an Unmanned Ground Vehicle in the Presence of Obstacles’, *IEEE Transactions on Automation Science and Engineering,* 18(4), pp. 1575–1589; https://doi.org/10.1109/TASE.2020.3010887.