

Ethical Challenges of AI-Enabled Autonomous Warfare

Eddie Henry II¹, Dike Illoh², Leonard Jones^{1*}, Brandon Mahoney¹, Winston Thompson¹

* Corresponding author: leonard.jones2@my.hamptonu.edu

¹Hampton University,
School of Engineering, Architecture, and Aviation
Hampton, VA

²University of Virginia,
School of Engineering and Applied Science
Charlottesville, VA

Abstract—In the last few decades, artificial intelligence has grown quickly, changing life for both civilians and military personnel. It has given them new abilities, but it has also raised a number of ethical and security issues. In the realm of autonomous warfare, AI systems augment military efficacy via sophisticated surveillance, data-informed targeting, and expedited decision-making processes. At the same time, autonomous systems also raise serious concerns related to accountability, escalation of conflict, system reliability, and unintended harm. This dual nature of AI’s development requires the need for an ethical examination of its role in modern warfare. This paper focuses on the ethical implications of AI-enabled autonomous warfare, with emphasis on human accountability, decision-making, transparency, and responsibility for harm in highly autonomous systems. The goal is to examine how the integration of AI into weapon systems challenges existing moral, legal, and institutional frameworks governing the use of force. The analysis focuses on autonomy in targeting and engagement decisions. This framework integrates principles from utilitarianism, just war theory, deontological ethics, and virtue ethics, including accountability and explainability. These frameworks are used to evaluate real-world and hypothetical scenarios involving autonomous military systems, such as automated targeting platforms and defensive response systems. The expected outcome is to have a clearer characterization of the ethical tensions created by autonomous AI in warfare, as well as to evaluate the conditions under which AI systems may support responsible military conduct. The results aim to inform ongoing debates on ethical guidelines and policy considerations for the development of AI-enabled autonomous weapons systems.

Keywords—Ethics, Artificial Intelligence, warfare, autonomous, Lethal Autonomous Weapon Systems

I. INTRODUCTION

Artificial intelligence (AI) is rapidly transforming modern warfare by enhancing military capabilities such as surveillance, decision-making, and autonomous operations. As AI technologies become more advanced, they are increasingly integrated into defense systems to improve efficiency, accuracy, and response time on the battlefield. [1] However, the growing use of AI in military applications also raises significant concerns about security, reliability, and global stability. Scholars warn that these technologies could shift the balance of power among nations, introduce new vulnerabilities, such as

cyberattacks, and potentially accelerate technological arms races [2].

Beyond strategic implications, the use of AI in warfare presents serious ethical and legal challenges. Autonomous weapon systems, which can identify and engage targets with limited human involvement, raise questions about accountability, decision-making, transparency, and the protection of civilian lives [3]. Existing international laws and ethical frameworks were designed around human decision-makers, making it difficult to determine responsibility when autonomous systems are involved. As AI continues to develop, it is essential to examine how these technologies can be regulated and ethically managed to ensure that military innovation does not undermine humanitarian principles or global security [4][5].

AI has achieved positive results in many aspects of warfare, including more effective targeting, increased accuracy and effectiveness of drones, and reducing overall costs of warfare due to increased efficiency. For example, Ukraine has rapidly adopted AI for first-person view (FPV) drones, loitering munitions, and autonomous navigation systems. These tools handle target recognition, tracking, last-mile guidance, and navigation (often bypassing GPS or jammed communications via onboard machine vision). FPV drone strike accuracy rose from 30–50% to around 80% in some units. Autonomous navigation increased success rates from 10–20% to 70–80%, reducing the need for multiple drones per target (e.g., from 8–9 to 1–2) and lowering overall costs and operator workload. Drones now account for roughly two-thirds of Russian losses and are described as twice as effective as other Ukrainian weapons. In Operation Spiderweb (June 2025), AI-assisted targeting helped damage 34% of Russia’s long-range bomber fleet, inflicting approximately \$7 billion in damage [6].

AI in warfare has also led to unnecessary death and destruction when used inappropriately. This occurs when the AI targets civilians and non-military infrastructure, and those targets aren’t properly vetted by humans. This is evident by the use of AI in the Israel-Gaza conflict. Since October 2023, the Israel Defense Forces (IDF) have used AI systems like Lavender (an AI-powered database for identifying suspected militants) and Gospel (for object/building targeting). Lavender processes vast amounts of data (e.g., from surveillance,

communications, and patterns of life) to assign scores to individuals and generate kill lists, often with minimal initial human review (as little as 20 seconds per target in early stages). Lavender reportedly flagged up to 37,000 Palestinians as potential Hamas/Islamic Jihad operatives (mostly junior), marking them and their homes for strikes. It accelerated targeting and airstrikes dramatically. The IDF has acknowledged using such tools for cross-referencing intelligence but maintains human oversight in final decisions. Reported accuracy was around 90% in samples, but the system permitted high collateral damage thresholds (e.g., 15–20 civilians per low-level target, over 100 for senior ones). This contributed to rapid operations but has been linked to significant civilian casualties and debates over proportionality and false positives (1 in 10 flagged individuals potentially misidentified per some accounts) [7].

These examples highlight the wide range of results possible with the use of AI in warfare. At its best, it will lower war costs, increase efficiency, and lower civilian casualties. At its worst, it will lead to massive death and destruction for any nation on the receiving end of AI-assisted warfare.

This also brings up an important conversation as to the arms race of AI technology. Similar to nuclear weapons, if AI-enabled warfare becomes the new frontier and the potential dangers aren't properly addressed and accounted for by a global coalition, the results can be disastrous.

II. ANALYZING AI-ENABLED AUTONOMOUS WARFARE THROUGH THE LENS OF VARIOUS ETHICAL FRAMEWORKS

AI-enabled autonomous warfare refers to Lethal Autonomous Weapon Systems (LAWS), which use AI to independently identify, select, and engage targets (including humans) without real-time human intervention. Examples include swarming drones, AI-piloted combat vehicles, or fully autonomous missile systems. The core ethical tension is the shift from human moral agency to algorithmic decision-making in life-and-death contexts. In this section, we will analyze this issue through various ethical frameworks, highlighting distinct concerns, trade-offs, and implications.

A. Utilitarianism

Utilitarianism evaluates actions by their outcomes. In other words, does this technology maximize overall well-being (or minimize suffering) for the greatest number? Benefits would include reducing friendly casualties by removing soldiers from the battlefield, enabling hyper-precise targeting to minimize collateral damage, and processing data faster than humans in chaotic environments. Supporters would argue that this will shorten conflicts, save civilian lives through better discrimination, and deter aggression via superior defensive capabilities. However, errors in AI perception, like misidentifying civilians as combatants due to training data biases or adversarial jamming, could cause mass unintended deaths. Long-term risks include an arms race, lowering the threshold for war due to cheaper and lower-risk conflicts, making leaders more willing to start one. This also opens up the possibility of proliferation to rogue actors or terrorists and

escalation into “flash wars” where AI systems react faster than humans can intervene. Utilitarians would demand rigorous empirical testing and cost-benefit modeling. If data shows net lives saved globally, deployment could be justified, but only with safeguards like “kill switches” and ongoing audits. Critics would argue that second-order effects, such as eroded global stability and AI unpredictability, would make a reliable calculation impossible.

B. Deontological Ethics

Deontological ethics and AI-enabled autonomous warfare lie at the crossroads of moral duties and the right to human life. Deontological ethics posits that humans have the moral duty to respect the liberties of others, an obligation that should be honored regardless of the situation. The Martens Clause serves as a deontological anchor in international law by requiring that technology be judged not by its “technical efficiency” but rather by the “dictates of public conscience” and “principles of humanity” [9], [11]. Morality in combat is not a computational problem of “optimization” but a duty-bound commitment to qualitative judgment that AI cannot replicate.

Delegating life-and-death decisions to an algorithm treats human beings as mere data points, thus violating the moral obligation to treat humanity as an end in itself. A key asset of deontology is the unconditional fulfillment of honoring the value of human lives, which cannot be done by a technological system. True agency requires the capacity for both empathy and compassion, qualities that are essential to ethical restraint and that cannot be imposed structurally by autonomous software [11]. Deontological ethics requires that human agency be maintained and control not fully placed in the hands of an AI system, as the duty to allow humans to control the lives of others is taken away in such systems.

Deontology also asserts that the value of human dignity is absolute and cannot be traded off for the utilitarian benefit of increased military precision or efficiency. Subjecting the loss of human life to an automated process is a direct challenge to the value of human dignity, as a person’s life is reduced to target characteristics processed by a model. Without the ability of a human decision-maker to make the moral choice of sparing a life based on extenuating circumstances, the value of a person as a human being is removed, which contradicts the basis of deontological ethics to uphold the value of human dignity and the right to life.

The utilization of AI-enabled autonomous warfare creates a larger gap between moral accountability and the consequences of a machine’s independent actions. Essentially, a “responsibility void” is created when an AI system is left to operate on its own without human intervention [8]. Justice for victims can potentially be rendered impossible by the existence of autonomous systems when there is no moral agent to experience guilt or face punishment for an unlawful attack. It is immoral to use the lives of others for personal benefit, and if there is no human present to experience this moral weight, the use of such warfare is seen as a deontological failure. Allowing machines to operate outside of a clear chain of human accountability undermines the professional and ethical standards of the military as a whole.

Despite the presentation of the linking conflicts between deontological ethics and AI-enabled autonomous warfare, there are counterarguments. Autonomous systems could theoretically be more accurate and fulfill the principle of distinction better than human soldiers, meaning that there is a “duty to deploy” such systems [11]. If these autonomous systems are not seen as emotionless instruments but rather as technological advancements that reduce the chance for human error, humans are obligated to utilize these systems, as they allow for the better preservation of human life. Deontological conflict arises between the duty to maintain human control and the competing duty to use the best available technology to minimize civilian casualties in unavoidable conflicts. Additionally, theories suggest that a machine’s strict adherence to programmed duties could prevent war crimes driven by human emotion. Removing the aspect of a human deviating from moral or ethical frameworks is seen as a helpful tool that automated systems could support.

A deontological assessment of autonomous warfare suggests that the moral credibility of lethal force cannot be separated from human agency and the ability to place value on human dignity. Technical advancements may offer greater precision, but they cannot single-handedly address the “accountability gap” and the reduction of human value to aggregated data points. Ultimately, the preservation of meaningful human control is a moral duty that ensures “public conscience” continues to govern war conduct. Any transition toward full autonomy in weapon systems must be weighed against the obligation to value human life and liberties with the gravity that only a moral agent can provide.

C. Just War Theory

Just War Theory provides a structured framework for evaluating the morality of warfare through two primary components: *jus ad bellum* (the justification for going to war) and *jus in bello* (the ethical conduct within war). In the context of AI-enabled autonomous warfare, most conversations are on the topic of *jus in bello*, particularly the principles of distinction, proportionality, and military necessity.

The main distinction requires combatants to be able to differentiate between military targets and civilians. Autonomous systems challenge this requirement because they rely on sensor inputs and training data, which may be incomplete, biased, or vulnerable to adversarial manipulation. While AI systems may offer improved pattern recognition and surveillance capabilities, failures in classification could result in civilians being incorrectly identified as legitimate targets. This raises concerns about whether autonomous systems can reliably meet the moral and legal standards required under international humanitarian law.

AI systems enhance assessments by analyzing large volumes of battlefield data. However, these calculations are based on programmed models that may not fully capture the moral weight of human life. Additionally, some AI systems complicate the ability to evaluate whether judgments are being made appropriately, raising concerns about accountability and transparency.

The principle of military necessity justifies the use of force only when it is required to achieve a legitimate military objective. Autonomous weapons may increase operational efficiency and speed, but their deployment in the field is a risk that could violate these principles.

A central issue across these principles is the concept of meaningful human control. Just War Theory assumes the presence of human moral agents who can be held accountable for decisions in warfare. When decision-making authority is delegated to autonomous systems, it becomes unclear who bears responsibility for violations of ethical or legal norms. This “accountability gap” challenges the application of Just War Theory, as responsibility may be diffused among developers, commanders, and political leaders rather than clearly assigned.

Scholars and policy organizations, including the International Committee of the Red Cross, emphasize that maintaining human oversight is critical to ensuring compliance with Just War principles. Without such oversight, AI-enabled systems risk undermining the ethical constraints that govern armed conflict. Therefore, from a Just War perspective, the use of autonomous weapons may only be justified if they can demonstrably uphold the principles of distinction and necessity while preserving meaningful human accountability. [8], [10]

D. Virtue Ethics

Virtue ethics, a framework derived from Aristotelian philosophy, focuses on an actor’s moral character and the development of good qualities rather than adherence to rules or evaluation of consequences. Instead of asking whether AI-enabled autonomous warfare produces desirable outcomes or follows moral laws, virtue ethics investigates how the use of such technologies influences the character, judgment, and moral responsibility of individuals and institutions involved. In this concept, prudence (practical knowledge), justice, accountability, and moral integrity are considered core qualities.

The delegation of life or death decisions to AI systems in the context of autonomous warfare raises serious questions regarding the diminution of human moral agency. Operators may become morally and psychologically detached from the results of their actions when they are removed from direct engagement. This disengagement runs the risk of undermining fundamental qualities like responsibility, compassion, and moral discernment. Such reliance on automation may eventually lead to “moral deskilling,” in which people stop actively practicing or developing the ethical reasoning needed in challenging and unpredictable combat situations [9].

Additionally, virtue ethics highlights the value of caution when making decisions, especially when there is uncertainty and moral ambiguity. Artificial intelligence (AI) systems are capable of processing enormous volumes of data and quickly and accurately identifying patterns, but they lack the moral reasoning and contextual awareness required for genuinely moral judgment. Making moral decisions in combat frequently necessitates weighing conflicting values, deciphering insufficient information, and using self-control skills that are still essentially human. By promoting blind faith in algorithmic

results, an over-reliance on AI runs the risk of undermining this important virtue [3].

Virtue ethics also extends beyond individuals to the character of institutions. The integration of autonomous weapons into military operations may gradually shape organizational values, potentially prioritizing efficiency, speed, and technological dominance over ethical reflection and human responsibility. If left unchecked, this shift could normalize a diminished role for human judgment in morally significant decisions, weakening the ethical culture within military institutions.

From a virtue ethics perspective, the ethical acceptability of AI-enabled autonomous warfare depends not only on the performance of the technology but on whether its use supports or undermines the development of virtuous moral agents. Maintaining meaningful human involvement in decision-making processes is essential to preserving accountability and moral responsibility. Autonomous systems should therefore be designed and implemented in ways that enhance, rather than replace, human ethical judgment. Ensuring that human operators remain actively engaged reinforces the cultivation of moral virtues and helps uphold the ethical standards necessary for responsible conduct in warfare [10] [12].

III. CONCLUSION

The integration of AI into autonomous warfare represents one of the most profound shifts in the nature of armed conflict in decades. As demonstrated through real-world applications such as Ukraine's effective use of AI-enhanced FPV drones and autonomous navigation systems like Spiderweb, which drastically improved strike accuracy and reduced operator workload. The use of AI-assisted targeting systems like Lavender in the Israel-Gaza conflict comes with concerns arising from high collateral damage thresholds and questions of proportionality, which have fueled debates over civilian harm. This technology offers tangible military advantages while simultaneously exposing deep vulnerabilities in human oversight and ethical judgment.

This paper has examined the ethical implications of AI-enabled autonomous warfare, with particular emphasis on autonomy in targeting and engagement decisions. By applying the lenses of utilitarianism, deontology ethics, Just War Theory, and virtue ethics, the analysis reveals persistent tensions. The potential for enhanced precision and reduced friendly casualties stands in contrast to risks of accountability gaps, diminished human moral agency, challenges to principles of distinction and proportionality, and the erosion of virtue ethics, such as prudence and responsibility. No one framework completely solves these problems. Utilitarian calculations hinge on uncertain net outcomes amid arms race dynamics and unpredictable errors. Deontological and Just War perspectives underscore the irreplaceable role of human moral judgment and the "responsibility void" created by full autonomy. Virtue ethics warns of long-term moral deskilling among operators and institutions.

At their core, these ethical frameworks converge on the imperative of meaningful human control. A concept that preserves human agency, ensures explainability when possible, and maintains the accountability chain in lethal decisions. While current AI systems can outperform humans in data processing and pattern recognition under controlled conditions, they lack contextual empathy, compassion, and nuanced ethical reasoning required for life-and-death choices in ambiguous battlefield environments. Giving such authority fully to algorithms risks reducing human lives to data points, undermining public conscience, and violating foundational principles of humanity embedded in international humanitarian law.

As AI capabilities continue to advance, the ethical examination presented here underscores that technological superiority must not come at the expense of moral integrity. Responsible military conduct in the AI era requires that innovation serves humanitarian principles, not replaces them. Future research and policy should focus on ensuring meaningful human control, developing verifiable, explainable AI for military applications, and fostering global cooperation to prevent an unregulated arms race. Only through such measures can AI-enabled systems support, rather than undermine, the responsible use of force in defense of shared human values.

ACKNOWLEDGMENT

Dr. Otsebele Nare - Electrical Engineering Department at Hampton University

Dr. Brian Aufderheide - Chemical Engineering Department at Hampton University

Dr. Garrick Louis - Department of Systems and Information Engineering at the University of Virginia

Thanks to research funding from the University of Virginia for conference registration

REFERENCES

- [1] C. Wilson, "Artificial Intelligence and Warfare," in 21st Century Prometheus: Managing CBRN Safety and Security Affected by Cutting-Edge Technologies, M. Martellini and R. Trapp, Eds. Cham, Switzerland: Springer, 2020. [Online]. Available: https://doi.org/10.1007/978-3-030-28285-1_7
- [2] N. S. Jecker, C. Atuire, V. Ravitsky, K. Behrens, and M. Ghaly, "War, Bioethics, and Public Health," *The American Journal of Bioethics*, vol. 25, no. 5, pp. 106–120, May 2025, <https://www.tandfonline.com/doi/epdf/10.1080/15265161.2024.2377118?needAccess=true>

- [3] N. G. Wood, "Explainable AI in the military domain," *Ethics and Information Technology*, vol. 26, no. 2, Art. no. 29, Apr. 2024, doi: 10.1007/s10676-024-09762-w.
- [4] Johnson, J. (2019). Artificial intelligence & future warfare: implications for international security. *Defense & Security Analysis*, 35(2), 147–169. <https://doi.org/10.1080/14751798.2019.1600800>
- [5] Simmons-Edler, Riley, et al. "Military AI Needs Technically-Informed Regulation to Safeguard AI Research and its Applications." arXiv, 2025, arXiv:2505.18371. arXiv, <https://arxiv.org/pdf/2505.18371>.
- [6] Kirichenko, D. (2025, August 21). *Artificial Intelligence's growing role in Modern Warfare*. War Room - U.S. Army War College. <https://warroom.armywarcollege.edu/articles/ais-growing-role/>
- [7] Abraham, Y. (2024, April 3). "Lavender": the AI Machine Directing Israel's Bombing Spree in Gaza. +972 Magazine. <https://www.972mag.com/lavender-ai-israeli-army-gaza/>
- [8] Lee, Jason. "Autonomous Weapons, War Crimes, and Accountability." *North Carolina Journal of International Law*, vol. 49, no. 2, 2024, [\https://journals.law.unc.edu/ncjil/wp-content/uploads/sites/3/2024/05/Autonomous-Weapons-War-Crimes-and-Accountability-by-Jason-Lee-24.pdf.\](https://journals.law.unc.edu/ncjil/wp-content/uploads/sites/3/2024/05/Autonomous-Weapons-War-Crimes-and-Accountability-by-Jason-Lee-24.pdf)
- [9] N. E. Sharkey, "The evitability of autonomous robot warfare," *International Review of the Red Cross*, vol. 94, no. 886, pp. 787–799, Jun. 2012, doi: <https://doi.org/10.1017/s1816383112000732>.
- [10] M. Horowitz and P. Scharre, "MEANINGFUL HUMAN CONTROL in WEAPON SYSTEMS: A Primer," Mar. 2015. Available: https://www.files.ethz.ch/isn/189786/Ethical_Autonomy_Working_Paper_031315.pdf
- [11] International Committee of the Red Cross (ICRC), "Autonomous Weapon Systems: Technical, Military, Legal and Humanitarian Aspects," 2014. <https://www.icrc.org/en/document/report-icrc-meeting-autonomous-weapon-systems-26-28-march-2014>
- [12] Stuart Russell, "Human Compatible: Artificial Intelligence and the Problem of Control." New York: Viking, 2019.