The Risk and Opportunity of Adversarial Example in Military Field

Yuwei Chen
Chinese Aeronautical Establishment
catcornic@gmail.com

Abstract

Artificial intelligence technology is increasingly widely used in the military field, and various countries have carried out a number of research and experiments, aiming to use artificial intelligence technology to shorten the closing time of their own kill chains, and obtain an advantage in the future battlefield, so as to increase the probability of victory in the battle. However, due to the vulnerability of deep learning models before adversarial examples, all systems or modules using artificial intelligence algorithms are at risk of being attacked, thereby delaying or hindering the closure of the opponent's kill chain and increasing the probability of combat victory from another aspect. Based on such risks, this paper proposes a conceptual scheme of military deception by attacking the AI modules of the combat units through adversarial examples, and proposes the challenges and prospects of the current technology. To the best of our knowledge, we are the first to analyze the impact of adversarial examples in the entire process of military operations, that is, the impact of each step and activity in the entire kill chain, and simulate the actual application of adversarial examples in combat through the wargame simulation platform. Ultimately, we found that when AI technology is really widely used in the military field, adversarial examples will have a subversive impact on several activities in several steps in the kill chain, which will directly lead to the interruption of the entire kill chain. This will lead to the failure of combat troops to successfully complete combat missions in accordance with the established objectives.

1. Introduction

In recent years, artificial intelligence algorithms based on deep learning has achieved many great achievements in many fields, profoundly affecting the development of society. In the fields of national defense, due to the perfect performance of AI algorithms, AI has increasingly become the core driving force to promote a new round of military revolution. At the same time, the future war with the characteristics of huge amount of data and difficult analysis puts forward more requirements on AI algorithms.

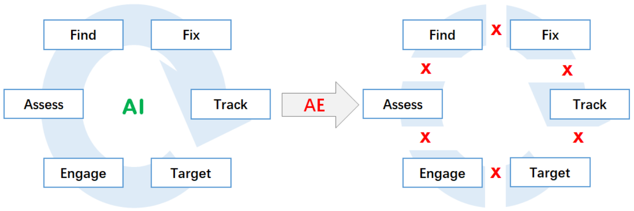

Figure 1. The left is the complete kill chain. The AI algorithms will greatly shorten the closed-loop time of the kill chain, but after the emergence of adversarial examples, the AI algorithms will be attacked, and the kill chain will not be able to form a closed loop normally just as the right.

However, the emergence of adversarial examples poses a great challenge to the security of AI technology applications. It is not difficult to see from various studies in recent years that adversarial examples can complete attacks on systems with AI algorithms in various scenarios, resulting in systems no longer capable of efficient and accurate analysis or judgment. It greatly reduces the credibility of AI algorithms in the application field. But at the same time, on the other hand, adversarial examples can be used as the object of algorithms learning, enhance the algorithm's understanding of the model mechanism, stimulate and make the AI algorithms more robust [27].

At present, some researchers have made preliminary attempts in the field of AI attack and defense in the military field using adversarial examples. In 2019, the U.S. Army Research Laboratory participated in a project using adversarial examples to attack AI recognition systems. In this project, researchers put stickers containing adversarial examples information on vehicles in the physical world, so that the vehicles cannot be accurately identified by commonly used detectors containing AI algorithms in urban and forest environments [28]. It poses a great challenge to the intelligent identification module in the future military satellite system. However, there is no work so far that can systematically and comprehensively analyze the effect of adversarial examples on military intelligent applications. Our contributions can be listed as follows:

- From the perspective of the military combat kill chain, we comprehensively analyze the challenges posed by adversarial examples in the military field to system modules that use artificial intelligence algorithms;
- For the typical six steps in the kill chain, we analyzed the possible impact of adversarial examples in each step;
- In the wargame simulation platform, we tested the subversive differences in the combat results when the adversarial examples participated.

2. Related Work

2.1. Adversarial Examples

In 2013, Google's Szegedy *et al.* proposed and defined adversarial examples that appear in the field of computer vision, a kind of tiny noise that is difficult to recognize by the human eyes, but can cause AI algorithms to produce wrong judgments [22].This brings great security risks to AI algorithms in the field of computer vision. Especially for the field of image recognition in AI algorithms, the attack of adversarial samples will cause a significant drop in the accuracy of deep learning [8, 16, 22, 24].

In addition, in different fields such as natural language processing [25] and audio recognition [4, 26], researchers have found that adversarial examples can be very aggressive to various types of AI algorithms and systems such as deep learning and reinforcement learning [2, 7, 9, 11, 14, 15].

More importantly, adversarial examples are not only effective in the virtual environment of the laboratory, but their attack effectiveness in the real physical world has also been gradually confirmed by researchers [3, 5, 17]. At present, adversarial examples attack many real-world fields including autonomous driving [6], face recognition [18], object detection [10], and robot navigation [16].

2.2. Application of Adversarial Examples in the Military Field

It can be seen from public academic materials that researchers have tried to introduce adversarial examples in the military field to achieve attacks on AI algorithms.

In 2018, Tomsett *et al.*, with support from the U.S. Army Research Laboratory and the UK Ministry of Defence, investigated the problems of machine learning model interpret-ability and susceptibility to adversarial examples and the impact they will have on future Army coalition operations equipped with AI technology [23].

In 2021, Haifeng Li *et al.* tried to attack SAR imaging function based on Convolutional Neural Network (CNN). They overlayed adversarial examples on the photos returned by the SAR imaging radars, successfully causing CNNs to identify tanks or other military units as false objects [13].

In addition, there are many projects targeting the intelligent modules of infrared recognition [20, 29] and image recognition [1, 12] in Intelligence, Surveillance, and Reconnaissance systems.

From these related studies, it is not difficult to see that most of the current research on the application of adversarial examples in the military field only stays in the detection activities in the operation, and does not analyze the application that may be involved in the whole process of the operation. Based on this, we carried out this research to conduct an in-depth analysis of the possibility and risks of adversarial examples in the full kill chain.

3. Opportunities and Risks of Adversarial Examples in the Military Field

In combat, the closing speed of the kill chain will directly affect the outcome of the war [19]. In the 6 major steps of the kill chain that called "Find-Fix-Track-Target-Engage-Assess", AI algorithms can help to quickly close the cycle of the kill chain.

$$\begin{aligned}\mathbf{T}_{KCLoop} = \mathbf{T}_{Find} + \mathbf{T}_{Fix} + \mathbf{T}_{Track} + \mathbf{T}_{Target} \\ + \mathbf{T}_{Engage} + \mathbf{T}_{Assess}\end{aligned}$$

The time of a kill chain is equal to the sum of the processing time of each step, and shortening the closure time of the kill chain means considering shortening the processing time of each step. Military researchers are embedding AI algorithms into materiels' systems that perform various steps of operations. For those steps that need to process huge amounts of data, AI algorithms can replace human warfighters and make accurate and rapid judgments, reducing the processing time of each step [21].

But after the concept of adversarial examples comes out, the kill chain will face disruptive challenges. According to the doctrine, the "Find, Fix, Track, Assess" steps tend to be ISR-intensive, while the "Target, Engage" steps are typically labor-, force-, and decision making- intensive [19].These are highly overlapping with the military problems that AI algorithms can solve, so adversarial examples can greatly impact the efficiency of AI algorithms in these steps, thus extending the closed-loop time of the kill chain that was originally expected to be reduced indefinitely.

3.1. Step 1: Find

In the "Find" step, the combat unit usually needs to analyze the received order, integrate the ISR resources that can be called at present, and assign the corresponding detection platforms to the designated area to collect battlefield situation information data according to the priority of the target in the order [19].

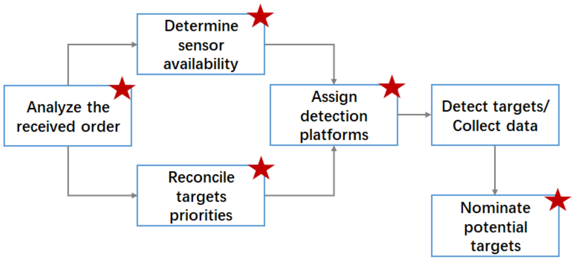
Determine sensor availability
Analyze the received order
Assign detection platforms
Detect targets/ Collect data
Reconcile targets priorities
Nominate potential targets

Figure 2. The main activities in the Find step. Activities marked with red five-pointed stars can use AI algorithms to improve the efficiency of activity processing.

For each activity in this step, AI algorithms can be embedded in platform systems such as satellites and aircraft to synthesize the huge amount of battlefield information and data, and analyze orders to achieve autonomous decision-making on areas to be detected, deployment of required resources, and nomination of potential targets, *etc*., makes various data collection and detection activities in the "Find" step more efficient, resulting in a reduction in $\mathbf{T}_{Find}$.

If the previous activity runs in the existing stable systems, and other modes such as cyberspace attacks are not considered, then the adversarial examples can only enter the system module that hosts the AI algorithm in the "Detect Targets/Collect Data" activity, and to attack the follow-up "Nominate potential targets" activity.

The attack may be in the form of: the adversarial examples are attached to the shape of the detected target, is collected by the system carrying out reconnaissance and detection activities, and the data is sent back to the AI analysis processing module, causing disturbance to the AI processing results and making it reach the wrong conclusion of resource calculation and allocation, resulting in insufficient resources to perform battlefield detection activities.

When there are insufficient resources to perform initial data collection and detection activities, the kill chain cannot obtain valid initial information to launch subsequent missions. This will greatly extend the $\mathbf{T}_{Find}$, which in turn affects the value of $\mathbf{T}_{KCLoop}$, and ultimately lead to an increase in the probability of losing the whole war.

3.2. Step 2: Fix

In this step, various sensors of the detection platforms will integrate terrain, astronomy and other environmental information of the potential target marked earlier to further locate and identify the target type, accurate locations and other information. Then, all kinds of target information is integrated and distributed to each operational unit according to the current available resources.

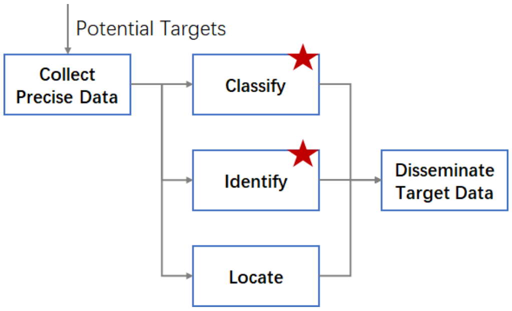

Figure 3. The main activities in the Fix step. Activities marked with red five-pointed stars can use AI algorithms to improve the efficiency of activity processing.

As a large amount of data needs to be fused in "Classify" and "Identify" activities, AI algorithms can replace pure manual to perform data integration and data analysis in these two activities, greatly improving the efficiency of target information acquisition and judgment.

In this step, the "Collect Precise Data" activity requires the sensor to collect information on the target. The adversarial examples can be received by the sensor at this time and transmitted to the following activities "Classify" and "Identify" as data to attack the AI algorithms existing in these two activities.

The form of attack may be consistent with that in the "Find" step. The adversarial examples can be attached to the key target in the form of appearance, material coating *etc*. After being collected, it will be transferred to the subsequent AI algorithm processing module, making the combat unit unable to accurately screen or identify the target type. As a result, the subsequent fusion of accurate target information is inconsistent with the actual situation, and an accurate operational plan cannot be generated, that is, $\mathbf{T}_{Fix}$ goes to infinity and so does $\mathbf{T}_{KCLoop}$. Finally,the kill chain cannot be closed.

3.3. Step 3: Track

In this step, the operational units will continuously track and monitor the identified targets while maintaining the collection of relevant situational information on the battlefield.

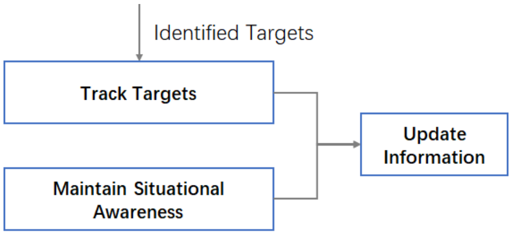

Figure 4. The main activities in the Track step.

The AI algorithms do not play a significant role in this step. However, for the adversarial examples, both "Track Targets" and "Maintain Situational Awareness" activities are generated by sensors to collect target information. The adversarial examples can be collected by sensors at this time. Although the adversarial example does not directly attack the activity in this step, once this information is added to the subsequent AI algorithm processing module along with the "Update Information" activity, it will most likely attack all subsequent AI algorithms.

3.4. Step 4: Target

"Target" is the most complex step in the entire kill chain, and it is also the stage of concentrated expression of the so-called art of war. It contains a lot of analysis and decision-making work, which is an important link between the previous and the next. This step usually needs to determine the target strike priority, evaluate the window of vulnerability, analyze the combat mission constraints such as weather and environment, determine the desired effectiveness of each target, evaluate the combat capability, estimate the collateral damage, confirm the rules of engagement, and maintain the continuous tracking. This step will eventually issue the final strike command order.

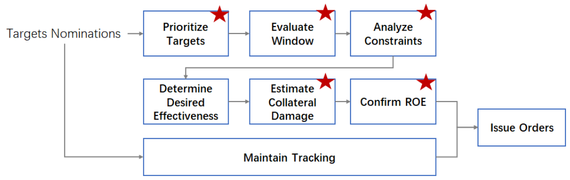

Figure 5. The main activities in the Target step. Activities marked with red five-pointed stars can use AI algorithms to improve the efficiency of activity processing.

Since the "Target" step needs to combine a huge amount of data and make judgments and analysis based on many constraints, the AI algorithms in this step can support activities including determining target priorities, evaluating strike windows, analyzing constraints, and evaluating collateral damage. This is to help commanders quickly generate final strike orders.

In this step, since the targets need to be tracked continuously, the sensors should continuously collect information on the targets, which allows the adversarial examples to have the opportunity to enter the systems. The attack of the adversarial examples in this step is often more fatal. It can target any AI algorithm module in this step according to the attacker's wishes, disrupting the overall decision-making of the entire mission, resulting in the final strike command order, and making the final order become invalid garbage.

3.5. Step 5: Engage and Step 6: Assess

In the "Engage" step and "Assess" step, what needs to be done is to execute the received orders, and evaluate the mission effectiveness according to the tasks. If the effect fails to achieve the expected mission effectiveness, it is necessary to determine whether to repeat attacks on targets or restart planning.

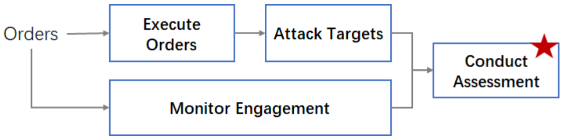

Figure 6. The main activities in the Engage and Assess steps. Activities marked with red five-pointed stars can use AI algorithms to improve the efficiency of activity processing.

In these two steps, AI algorithms can be used as analytical means of evaluation to assist commanders in evaluating combat effectiveness.

At this time, what the adversarial example can do is to pass the information to the enemy's AI algorithm module through the interaction during the battle, and guide the attack control module or the mission effectiveness evaluation module to make a wrong judgment, which not only makes the attack activity unable to be executed normally, but also causes commanders cannot make the right decisions based on that assessment information. So eventually, the $\mathbf{T}_{Assess}$ will be extended.

If the adversarial examples effectively attack the AI algorithms in each of the above steps, the total kill chain time $\mathbf{T}_{KCLoop}$ will become as follows.

$$\begin{aligned}\mathbf{T}'_{KCLoop} = \mathbf{T}'_{Find} + \mathbf{T}'_{Fix} + \mathbf{T}_{Track} + \mathbf{T}'_{Target} \\ + \mathbf{T}_{Engage} + \mathbf{T}'_{Assess}\end{aligned}$$

where

$$\begin{aligned}\mathbf{T}'_{Find} > \mathbf{T}_{Find} \qquad \mathbf{T}'_{Fix} > \mathbf{T}_{Fix} \\ \mathbf{T}'_{Target} > \mathbf{T}_{Target} \qquad \mathbf{T}'_{Assess} > \mathbf{T}_{Assess}\end{aligned}$$

So as a result, $\mathbf{T}'_{KCLoop}$ is significantly greater than $\mathbf{T}_{KCLoop}$. It means that adding adversarial examples during military missions will effectively prolong the closing time of the kill chain, thereby reducing combat efficiency and increasing the possibility of war failure.

4. Experiments and Analysis

To test the real impact of adversarial examples on the kill chain in combat scenarios, we take an offensive counter-air mission in a certain area as an example to analyze how subversive changes the adversarial examples can bring to military operations.

4.1. Experimental Setup

Using the “Command Modern Operations” software, we set red and blue against both sides, with the red side being the defending side and the blue side being the attacking side. Blue’s specific combat units include reconnaissance satellites, RQ-180 unmanned reconnaissance aircraft, E-3 early warning aircraft, and F-35 fighters. The specific units of the red side participating in the battle are airport ground facilities and F-22 fighters. The main combat mission of the blue side is to attack the airfield targets of the red side.

The specific combat tasks are set as follows: the blue side dispatches the RQ-180 reconnaissance aircraft to conduct reconnaissance in the area where the target is located, and after determining the target position information, the F-35 fighters are dispatched to attack the target, and the E-3 early warning aircraft provides battlefield awareness for the F-35 fighters. When the red side finds the incoming enemy plane, it will direct the air defense system to attack the air hostile target, and at the same time send F-22 fighters to the target area to clear the airspace.

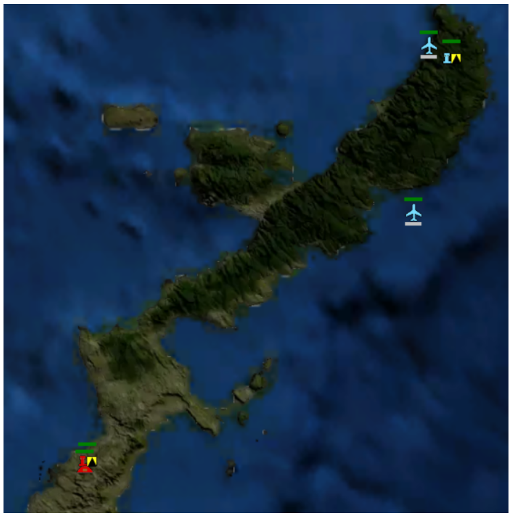

Figure 7. The initial situation of the red and blue sides

In this scenario, adversarial examples can be set up to attach to the red side airfield targets. Assuming that the coefficient ϵ by which the target can be fully detected and identified by the reconnaissance unit is 1, when the target is loaded with an adversarial example, the coefficient ϵ is set to a random value in the range of 0.1 to 0.9 in the software. In the experiment, through the detection time $\mathbf{T}_d$, the location time $\mathbf{T}_l$, the ammunition loss $\mathbf{E}$, the number of aircraft losses $\mathbf{L}$, the total task completion time $\mathbf{T}_{total}$ and other indicators, the completion of the task when the adversarial examples are not loaded and when the adversarial examples are loaded are respectively tested.

4.2. Engagement without adversarial examples

In the absence of adversarial examples, the blue side’s RQ-180 reconnaissance aircraft found the target almost immediately after adjusting the detection direction, completed the precise positioning, and started the stable tracking mode at the same time.

The blue side’s F-35 fighter took off under the guidance of the command center, and headed to the scheduled bombing site according to the target information returned by the RQ-180 reconnaissance plane.

When the red side’s air defense system detects an enemy target in the air, it directs the air defense system to automatically intercept the incoming enemy target. At the same time, the F-22 fighters take off and go to the incoming area of the enemy target to clear the airspace.

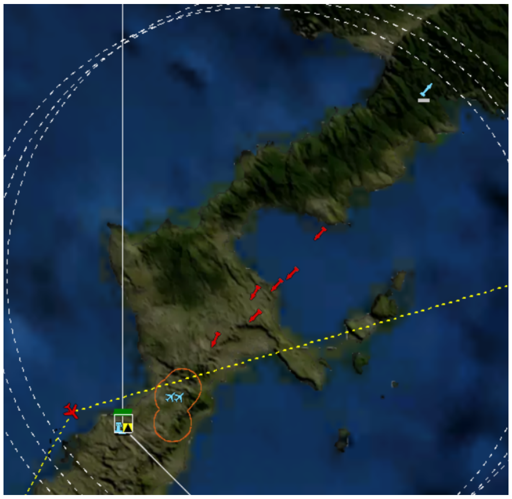

Figure 8. Engagement without adversarial examples, normal red and blue battlefield situation.

The battle ended with the blue side’s F-35 fighters successfully taking out the target, which in this case took just 6 minutes.

The experimental results are as follows: where ϵ_1 is 1, the detection time T_{d1} is 1 second, the location time T_{l1} is 1 second, the ammunition loss E_1 is 6, the number of aircraft losses L_1 is 0, and the total task completion time T_{total1} is about 6 minutes.

4.3. Engagement with Adversarial Examples

In this scenario, the adversarial examples are attached to the surface of the red side’s target by means of textures, spotlights, *etc*., and are simulated in the simulation system by changing the target characteristics of the red side’s targets.

At the beginning of the mission, after adjusting the detection orientation, the blue side’s RQ-180 reconnaissance aircraft was able to scan the buildings in the target area, and sent the reconnaissance information back to the artificial intelligence module. As a result, the building could not be identified as a target to be attack, and continuous reconnaissance of the target was abandoned.

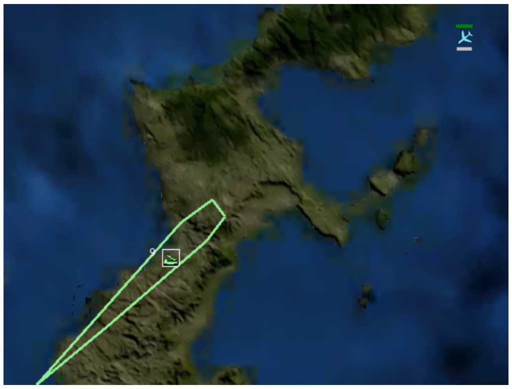

Figure 9. Engagement with adversarial examples, the blue side's reconnaissance aircraft cannot accurately identify and judge the target information.

At this time, the blue side's commander could only judge that there was indeed an attack target in the area based on the intelligence, so he sent the RQ-180 reconnaissance aircraft to continue to conduct reconnaissance task in the area.

When the RQ-180 reconnaissance aircraft approached the core area, it received the reconnaissance information containing the adversarial examples again, and sent it back to the artificial intelligence module for processing. The only way is to continue to approach it to carry out reconnaissance, which further compresses the original effective detection range of the RQ-180 reconnaissance aircraft. At this time, the blue side's commander still could not obtain accurate information on the target, and could not carry out further operational order planning.

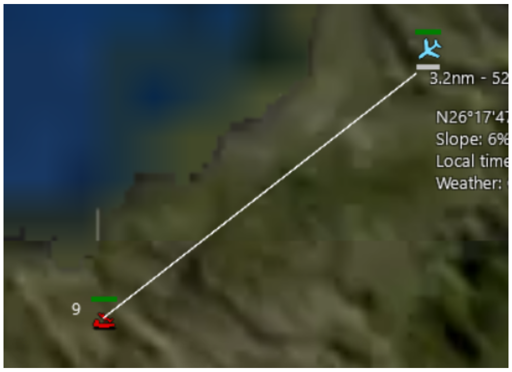

Figure 10. Engagement with adversarial examples, The blue side's reconnaissance aircraft had to get close to the target in order to collect more precise target information.

When it is about 3 nautical miles away from the final target, the RQ-180 reconnaissance aircraft realizes the identification and stable tracking of the target. But at this time, because it was too close to the red side's basement, the reconnaissance aircraft was quickly shot down by the red side's air defense systems, and the blue side's offensive intention was seen by the red side.

The red side quickly turned on the offensive and defensive mode, which made the blue side's probability of its successful attack fall off a cliff, forcing the blue commander to abandon the attack mission, and the blue side's mission failed.

The experimental results are as follows: where $\epsilon_2 \in [0.1, 0.9]$, the detection time T_{d2} is 358 seconds, the location time T_{l2} is 705 seconds, the ammunition loss E_2 is 16, the number of aircraft losses L_2 is 4, and the total task completion time T_{total2} is about 31 minutes.

4.4. Comparison of Simulation Results

Comparing the results of the two scenarios, it is not difficult to see that even using adversarial examples to attack artificial intelligence algorithms in the "Find" and "Fix" steps still has an important impact on the overall military operations.

After adding the adversarial examples to the target characteristics, the red side did not suffer large asset losses in the simulation game, and even the ammunition of the air defense system did not consume too much. That is to say, they won the defense and predicted the blue side's combat intention. As a result, the blue side has changed from taking the initiative to a very passive situation.

5. Conclusions

When adversarial examples appear in combat scenarios, the artificial intelligence algorithms that can be deeply relied on in the past will instantly become the most unreliable link, and the next steps of the enemy and the enemy will become more obscure. In this research, the possible application of adversarial examples at various steps in the kill chain is analyzed, and at the same time, we verify the lethal impact of adversarial examples on combat missions in a virtual simulation environment.

It is not difficult to predict that this situation will in turn slow the deployment of AI in the military field. Military experts will use AI in a wide range with artificial means after the AI modules are sufficiently stable and robust. How to improve the security of one's own AI and how to use adversarial examples to reduce the reliability of the enemy's AI will become the focus of subsequent research on AI in the military field.

References

[1] Ajaya Adhikari, Richard den Hollander, Ioannis Tolios, Michael van Bekkum, Anneloes Bal, Stijn Hendriks, Maarten Kruithof, Dennis Gross, Nils Jansen, Guillermo Pérez, Kit Buurman, and Stephan Raaijmakers. 2

[2] Vahid Behzadan and Arslan Munir. Vulnerability of deep reinforcement learning to policy induction attacks. In *International Conference on Machine Learning and Data Mining in Pattern Recognition*, pages 262–275. Springer, 2017. 2

[3] Tom B Brown, Dandelion Mané, Aurko Roy, Martín Abadi, and Justin Gilmer. Adversarial patch. *arXiv preprint arXiv:1712.09665*, 2017. 2

[4] Nicholas Carlini and David A. Wagner. Audio adversarial examples: Targeted attacks on speech-to-text. *CoRR*, abs/1801.01944, 2018. 2

[5] Ranjie Duan, Xingjun Ma, Yisen Wang, James Bailey, A Kai Qin, and Yun Yang. Adversarial camouflage: Hiding physical-world attacks with natural styles. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1000–1008, 2020. 2

[6] Kevin Eykholt, Ivan Evtimov, Earlence Fernandes, Bo Li, Amir Rahmati, Chaowei Xiao, Atul Prakash, Tadayoshi Kohno, and Dawn Song. Robust physical-world attacks on deep learning visual classification. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1625–1634, 2018. 2

[7] Adam Gleave, Michael Dennis, Cody Wild, Neel Kant, Sergey Levine, and Stuart Russell. Adversarial policies: Attacking deep reinforcement learning. *arXiv preprint arXiv:1905.10615*, 2019. 2

[8] Ian J Goodfellow, Jonathon Shlens, and Christian Szegedy. Explaining and harnessing adversarial examples. *arXiv preprint arXiv:1412.6572*, 2014. 2

[9] Yi Han, Benjamin IP Rubinstein, Tamas Abraham, Tansu Alpcan, Olivier De Vel, Sarah Erfani, David Hubczenko, Christopher Leckie, and Paul Montague. Reinforcement learning for autonomous defence in software-defined networking. In *International Conference on Decision and Game Theory for Security*, pages 145–165. Springer, 2018. 2

[10] Lifeng Huang, Chengying Gao, Yuyin Zhou, Cihang Xie, Alan L. Yuille, Changqing Zou, and Ning Liu. Universal physical camouflage attacks on object detectors. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 717–726, 2020. 2

[11] Sandy Huang, Nicolas Papernot, Ian Goodfellow, Yan Duan, and Pieter Abbeel. Adversarial attacks on neural network policies. *arXiv preprint arXiv:1702.02284*, 2017. 2

[12] Hyun Kwon, Yongchul Kim, Hyunsoo Yoon, and Daeseon Choi. Fooling a neural network in military environments: Random untargeted adversarial example. In *MILCOM 2018 - 2018 IEEE Military Communications Conference (MILCOM)*, pages 456–461, 2018. 2

[13] Haifeng Li, Haikuo Huang, Li Chen, Jian Peng, Haozhe Huang, Zhenqi Cui, Xiaoming Mei, and Guohua Wu. Adversarial examples for cnn-based sar image classification: An experience study. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14:1333–1347, 2021. 2

[14] Jieyu Lin, Kristina Dzeparoska, Sai Qian Zhang, Alberto Leon-Garcia, and Nicolas Papernot. On the robustness of cooperative multi-agent reinforcement learning. In *2020 IEEE Security and Privacy Workshops (SPW)*, pages 62–68. IEEE, 2020. 2

[15] Yen-Chen Lin, Zhang-Wei Hong, Yuan-Hong Liao, Meng-Li Shih, Ming-Yu Liu, and Min Sun. Tactics of adversarial attack on deep reinforcement learning agents. *arXiv preprint arXiv:1703.06748*, 2017. 2

[16] Aishan Liu, Tairan Huang, Xianglong Liu, Yitao Xu, Yuqing Ma, Xinyun Chen, Stephen Maybank, and Dacheng Tao. Spatiotemporal attacks for embodied agents. In *European Conference on Computer Vision*, 2020. 2

[17] Aishan Liu, Xianglong Liu, Jiaxin Fan, Yuqing Ma, Anlan Zhang, Huiyuan Xie, and Dacheng Tao. Perceptual-sensitive gan for generating adversarial patches. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pages 1028–1035, 2019. 2

[18] Luan Nguyen, Sunpreet S. Arora, Yuhang Wu, and Hao Yang. Adversarial light projection attacks on face recognition systems: A feasibility study. *CoRR*, abs/2003.11145, 2020. 2

[19] Joint Chiefs of Staff. Joint publication 3-60 joint targeting. 2, 3

[20] Anthony Ortiz, Olac Fuentes, Dalton Rosario, and Christopher Kiekintveld. On the defense against adversarial examples beyond the visible spectrum. In *MILCOM 2018 - 2018 IEEE Military Communications Conference (MILCOM)*, pages 1–5, 2018. 2

[21] Peter Svenmarck, Linus Luotsinen, Mattias Nilsson, and Johan Schubert. Possibilities and challenges for artificial intelligence in military applications. In *Proceedings of the NATO Big Data and Artificial Intelligence for Military Decision Making Specialists' Meeting*, pages 1–16. Neuilly-sur-Seine France, 2018. 2

[22] Christian Szegedy, Wojciech Zaremba, Ilya Sutskever, Joan Bruna, Dumitru Erhan, Ian Goodfellow, and Rob Fergus. Intriguing properties of neural networks. 12 2013. 2

[23] Richard Tomsett, Amy Widdicombe, Tianwei Xing, Supriyo Chakraborty, Simon Julier, Prudhvi Gurram, Raghuveer Rao, and Mani Srivastava. Why the failure? how adversarial examples can provide insights for interpretable machine learning. In *2018 21st International Conference on Information Fusion (FUSION)*, pages 838–845, 2018. 2

[24] Kun Wei, Muli Yang, Hao Wang, Cheng Deng, and Xianglong Liu. Adversarial fine-grained composition learning for unseen attribute-object recognition. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 3741–3749, 2019. 2

[25] Cihang Xie, Jianyu Wang, Zhishuai Zhang, Yuyin Zhou, Lingxi Xie, and Alan Yuille. Adversarial examples for semantic segmentation and object detection. 03 2017. 2

[26] Xuejing Yuan, Yuxuan Chen, Yue Zhao, Yunhui Long, Xiaokang Liu, Kai Chen, Shengzhi Zhang, Heqing Huang, Xiaofeng Wang, and Carl A Gunter. {CommanderSong}: A

systematic approach for practical adversarial voice recognition. In *27th USENIX security symposium (USENIX security 18)*, pages 49–64, 2018. 2

[27] Chongzhi Zhang, Aishan Liu, Xianglong Liu, Yitao Xu, Hang Yu, Yuqing Ma, and Tianlin Li. Interpreting and improving adversarial robustness with neuron sensitivity. *IEEE Transactions on Image Processing*, 2020. 1

[28] Yang Zhang, Hassan Foroosh, Philip David, and Boqing Gong. CAMOU: Learning physical vehicle camouflages to adversarially attack detectors in the wild. In *International Conference on Learning Representations*, 2019. 1

[29] Xiaopei Zhu, Xiao Li, Jianmin Li, Zheyao Wang, and Xiaolin Hu. Fooling thermal infrared pedestrian detectors in real world using small bulbs. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 3616–3624, 2021. 2