

인공지능과 로봇의 군사적 활용과 선행 과제

Military Use of Artificial Intelligence and Robots and the Prerequisite

송윤선¹⁾

Yun-seon Song

[초 록]

인공지능, IoT, 빅데이터, 로봇, 자율주행차량, 드론 등 4차산업혁명 신기술의 급격한 발전으로 인해 미래전 수행방식은 지금과는 완전히 다른 모습을 띠게 될 것이다. 인공지능과 로봇은 다른 기술들과 융합하여 완전히 새로운 개념의 자율 무기체계로 발전할 것이다. 인간이 주체가 된 기존의 전쟁과 달리 자율무기 체계가 전쟁 행위의 주체로서 참여하게 된다면, 정치적, 법적, 윤리적 문제를 비롯한 우리가 예상하지 못한 여러 가지 문제가 발생할 수 있다. 이러한 점에서 인공지능과 로봇의 군사적 응용에 대한 이해와 이로 인해 발생할 수 있는 새로운 문제들을 식별하여 선행적으로 대비해야 할 방향을 제시하고자 한다.

[ABSTRACT]

New technology of the 4th Industrial Revolution, such as artificial intelligence, IoT, big data, robots, and so on, will drive the warfare in the future into a new aspect. AI and robot technology will produce a new concept of autonomous weapon systems by the fusion with other relevant technology. If autonomous military machinery are to take part in wars as a principal actor, differently with previous wars in which the human have played as main actors, there may be many problems that we don't expect, including political, legal, and ethical problems. In this regard, I will explain the application of AI and robot technology in the military, identify the new issues that may be produced subsequently and present some policy suggestions for preparing a future war waged by autonomous machinery.

Key Words : 인공지능(AI, Artificial Intelligence), 자율 무기체계(Autonomous Weapon System)

1. 서론

최근 4차산업혁명이라 불리는 신기술들의 급격한 발전은 인류 사회의 모든 분야에 큰 영향을 미치며 인간의 삶을 새로운 양상으로 이끌어가고 있다. 여기에는 군사 분야도 예외는 아니다. 미래전 수행방식은 지금과는 완전히 다른 모습을 띠 것이라고 대부분의 미래학자와 군사전문가들은 전망하고 있다. 이러한 전망은 크게 빗나가지 않을 것이다. 전쟁의 역사와 과학기술이 밀접한 관계를 가지고 있음을 말해 준다. 1차산업혁명을 통해 기계가 출현하면서 인간의 완력에 의존하던 전쟁이 기계에 의존하는 전쟁으로 변환하였으며, 이어서 1950년대 이후 컴퓨터와 통신 기술이 발전하면서 네트워크 중심의 전

쟁으로 진화하였다. 앞으로 초지능, 초연결 시대의 전쟁은 어떠한 모습을 띠 것인가? 미래전 수행 방식에 가장 큰 영향을 미치게 될 기술은 인공지능과 로봇일 것이다. 이것은 인간의 삶에 영향을 가장 크게 미칠 기술과도 크게 다르지 않다. 전쟁 역시 인간 삶의 일부이기 때문이다. 인공지능은 연관 기술들과 결합하여 인간의 두뇌 활동을 대신할 것이며, 로봇은 인간의 육체 활동을 대신할 것이다.

아직은 초기 단계이기는 하지만 인공지능과 로봇이 군사 분야에 도입되어 그 활용범위가 확대되고 있다.²⁾ 앞으로 인공지능과 로봇은 그 기술 수준이 높아지면서 다른 기술들과 융합하여 완전히 새로운 개념의 자율 무기체계로 발전할 것이다. 그

²⁾ 사실 재래식 무기체계에도 자동화 또는 무인 기능이 부분적으로 적용되어왔다. 예를 들어 현재 한국군에서 십여 년 전부터 운용하고 있는 GOP 경계 무인화 시스템도 일종의 부분적인 자율시스템이라고 할 수 있으며, 미사일 자동 교전 시스템 등도 큰 의미의 로봇화 시스템으로 볼 수 있다. 그러나 본 고에서 말하는 인공지능이나 로봇은 초연결, 초지능에 기반을 둔 신개념적 차원의 기술을 일컫는다.

1) 국민대학교 정치대학원 (The Graduate School of Politics and Leadership, Kookmin University, Korea)

E-mail: sys70700@gmail.com

Copyright © The Korean Institute of Defense Technology

Received : May 31, 2021

Revised :

Accepted : June 17, 2021

결과 미래전은 인간이 아닌 기계들의 전쟁이 될 것이라고 많은 사람들이 전망한다. 이러한 극단적인 모습은 아닐지라도 머지않은 미래의 전쟁은 인간 전투원과 지능형 군사로봇과의 협업으로 진행될 것이 분명해 보인다. 이러듯, 인간이 주체가 된 기존의 전쟁과 달리 자율무기 체계가 전쟁 행위의 주체로서 참여하게 된다면, 정치적, 법적, 윤리적 문제를 비롯한 우리가 예상하지 못한 새로운 여러 가지 문제가 발생할 수 있다. 이러한 점에서 이 글은 인공지능과 로봇 기술이 중심이 된 자율 무기 체계의 군사적 활용에 대한 이해와 이로 인해 발생할 수 있는 새로운 문제들을 식별하고, 우리 군이 선행적으로 대비해야 할 방향을 제시하고자 한다.

2. 인공지능과 로봇 기술의 군사적 적용

2.1 인공지능과 로봇의 개념

인공지능이란 사고나 학습 등 인간이 가진 지적 능력을 컴퓨터를 통해 구현하는 기술이다. 인공지능은 개념적으로 바둑을 두는 알파고처럼 특정 분야에 특화된 형태로 개발되어 인간의 한계를 보완하고 생산성을 높이기 위해 활용되는 약인공지능(Weak AI)과 인간처럼 여러 분야에 적용할 수 있으며 지각과 사고가 가능한 강 인공지능(Strong AI) 또는 범용 인공지능(AGI)으로 구분할 수 있다. 사실 인공지능의 개념이 등장한 것은 1950년대이지만, 오랫동안 답보상태에 있다가 반도체 기술의 발전에 힘입어 2000년대 초반 기술적 도약을 시작하여 현재 4차산업혁명을 견인하는 대표적인 기술이다. 현재까지 개발된 인공지능은 모두 약인공지능이며, 강인공지능은 아직 등장하지 않았지만 머지않아 개발될 것으로 전망된다.³⁾

한편, 로봇은 어원상 힘든 일이나 강제적인 노역을 인간 대신 수행하는 기계를 의미하는 용어로서,⁴⁾ 국제로봇연맹(International Federation of Robotics)의 정의에 따르면, '외부의 환경을 자체적으로 인식(perception)하고, 스스로 그 상황을 판단(cognition)하여 자율적으로 동작(mobility & manipulation)하며 주어진 임무를 수행하는 기계'를 뜻한다. 로봇은 임무와 용도에 따라 용접, 핸들링, 도장 등 생산 현장에서 산업 자동화용으로 사용되는 산업용 로봇과 청소, 의료, 구조, 노약자 보조 등 인간과 설비에 유용한 서비스를 제공하는 서비스 로봇으로 구분하고 있다. 최근에 많이 불리는 지능형 로봇은 인공지능을 탑재하여 첨단기능을 보유한 로봇을 포괄적으로 의미한다. 따라서 인공지능과 로봇은 상호 밀접한 관계를 가지고 있으며, 이것이 상호 융합되었을 때 인간의 삶에 파격적 변혁을 불러오게 된다.

3) Ray Kurzweil은 강인공지능이 2045년경에 등장할 것으로 예측하며, 이 시기를 '기술적 특이점(Technological Singularity)'으로 부르고 있다. Ray Kurzweil, 『Singularity is near』 (Penguin Books, 2006).

4) 로봇(robot)이란 용어는 체코슬로바키아의 소설가 차페크(Karel Capek)가 1921년 발간한 『R.U.R(Rossum's Universal Robots)』이라는 희곡에서 처음으로 사용되었다. 로봇의 어원이 체코어의 노동을 의미하는 단어 '로보타(robota)'인 만큼, 로봇의 역할은 인간의 노동을 대신 수행하는 데서 찾을 수 있다. <https://100.daum.net/encyclopedia/view/14XXE0074082> (검색일: 2021.7.16.)

2.2 인공지능과 로봇 기술의 군사적 활용

인공지능과 로봇 기술은 군사 분야에 지대한 영향을 미치고 미래전 수행방식에 큰 변혁을 일으킬 것이다. 인공지능 기술은 다른 연관 기술들과 융합하여 다양한 기능을 수행할 수 있다. 인공지능 기술이 빅데이터와 클라우드 기술과 융합하여 센서에서 수집된 데이터를 통해 정보분석, 적 능력 및 기도를 판단하거나, 위계임과 전장 시뮬레이션을 통해 방책을 분석 및 평가하여 최적의 대안을 제시할 수 있다. 또한 표적 성질에 따른 최적의 표적 할당과 아군의 상황을 고려한 최적의 전투 편성 방안을 도출해 낼 수 있으며, 스마트 그리드 기술⁵⁾을 응용하여 전투 물자와 장비 분배를 최적화할 수 있다. 또한 작전계획 수립, 명령 작성 및 하달 등 일련의 작전 지원 활동을 인간의 감독하에 자동으로 신속 정확하게 처리할 수 있다. 특히 디지털 영역인 사이버 작전에 있어 인공지능 기술의 위력이 최대로 발휘될 수 있다. 인공지능에 의해 수행되는 컴퓨터 네트워크의 프로빙(probing), 매핑(mapping), 해킹(hacking) 기술을 활용하여 적의 사이버 취약점을 신속하게 파악하여 공격하거나 아군의 취약점을 식별하고 이를 보강하는 등 사이버전 수행 능력을 획기적으로 향상할 수 있다. 이러한 인공지능의 기능들은 지금까지 인간이 직접 수행하던 것들로서 인간의 인지적 한계와 전문인력의 제한, 시간 지연 등으로 인해 어쩔 수 없이 발생하는 문제들을 일거에 해소하고 완전 작전을 수행할 수 있도록 해줄 것이다.

이러듯 인공지능이 ICT 기술과 결합하여 인지와 의사소통 등 인간의 소프트 영역을 대체하는 한편, 인공지능이 기계와 결합하여 지능형 군사로봇 또는 자율 무기체계로서 인간의 물리적 영역을 획기적으로 대체할 것이다. 지능형 군사로봇은 지능형 로봇이 가지고 있는 이동성과 지능을 포함하여, 군인이 수행하는 임무나 기존에 불가능했던 새로운 임무를 무인 자율 또는 원격제어를 통해 수행할 수 있는 로봇 시스템으로서, 현재에도 군에서 초보적 수준에서 다양하게 활용되고 있으며, 점점 그 범위가 확대되고 있다. 공중에서는 무인항공기 및 UAV, 드론, 조류형 로봇, 지상에서는 동물형 로봇, 차륜형 및 궤도형 로봇, 해양에서는 물고기 로봇, 무인잠수정, 무인수상함미 개발되어 현장에서 활동하고 있거나 현재 개발 중에 있다.

2.3 자율 무기체계의 개념과 특성

인공지능과 로봇 기술의 군사적 결합은 자율 무기체계의 형태로 나타날 것이다. 인공지능은 자율 무기체계의 두뇌를 구성하며, 로봇은 곧 몸통이다. '자율(autonomy)'이라는 것은 자신의 행동을 스스로 통제하고 책임을 지는 것을 말하는 것으로, 이는 무기에 대한 인간의 통제가 사라지고 무기 자신이 스스로 통제함을 뜻한다. 지금까지 무기는 대부분 인간의 직접적인 작동 또는 통제로 운용되는 수동적인 전투 수단이었으나, 자율무기는 이러한 기존의 무기체계와는 완전히 다른 개념으로 무기 체계가 인간 전투원과 같이 하나의 독립된 전투 주체가 된다. 그러나 현실적으로는 공상 과학 영화에서 보는 것과 같은 완전

5) 기존의 전력망에 정보통신기술(ICT)을 접목해 전력 공급자와 소비자가 양방향으로 실시간 전력 정보를 교환함으로써 에너지 효율을 최적화하는 차세대 지능형 전력망을 뜻한다.

한 자율성을 가진 무기체계는 아직 개발되지 않았다. 다만, 자동화 시스템에 의해 무기체계가 스스로 표적을 탐지하면 인간의 결정에 따라 발사되어 목표를 자동으로 추적하여 타격하는 ‘반자율적(semi-autonomous)’ 수준에 이르고 있다. 만일 무기체계가 인간의 개입 없이 스스로 목표를 선정하고 공격 여부를 결정한다면 이는 완전한 형태의 자율무기라고 할 수 있다.

자율성(autonomy)의 개념을 좀 더 명확히 이해하기 위해서는 ‘자동화(automation)’의 개념과 비교하는 것이 도움이 된다. 자동화는 센서를 통한 외부의 데이터 입력에 반응하여 미리 정해진 절차에 따라 작동하는 것을 의미한다. 자율성 또한 정해진 절차에 따라 작동하는 것을 의미하지만 양자의 결정적인 차이는 알고리즘 작동이 폐쇄적이냐 개방적이냐이다. 즉, 자율성이 자동화와 다른 점은 기계학습을 통해 환경의 변이를 인식하고 스스로 이에 적응할 수 있다는 것이다.⁶⁾ 자동화와 자율성의 이러한 구분은 그러나 현실적으로 두 가지가 겹쳐서 나타나는 중간지대가 많아 명확히 구분하기 어려운 실정이다. 다만, 자율성의 정도를 기계에 대해 인간이 관여하는 정도에 따라 구분해 볼 수 있다. 자율성의 구분은 미국 국방부의 ‘무기체계의 자율성에 관한 지침’이 널리 준용되고 있다.⁷⁾ 미 국방부는 전투 행동의 순환 과정인 관측(observe) - 사고(orient) - 판단(decide) - 행동(act)의 고리를 의미하는 OODA Loop에 인간이 어느 정도 개입과 통제를 하는가에 따라 자율성을 <표 1>과 같이 ‘human in the loop’, ‘human on the loop’, human ‘out of the loop’의 세 가지 범주로 구분하였다.

구분	내용
1단계 (human in the loop)	가장 낮은 수준의 자율성으로, 인간이 OODA loop 전 과정에서 무기체계의 작동에 관여하는 상태
2단계 (human on the loop)	중간 수준의 자율성으로, 무기체계가 독립적으로 작동하나 인간의 감독하에 이루어지는 상태
3단계 (human out of the loop)	완전한 수준의 자율성으로, OODA loop 전 과정에서 인간의 개입 없이 기계 스스로 독립적으로 작동하는 상태

표 1. 무기체계의 자율성 범주 구분

Table 1. The category of autonomy of weapon system

1단계 ‘human in the loop’는 인간 운용자가 전투 행위의 전 과정에 걸쳐 무기체계를 직접 개입하여 통제하는 것으로 원격으로 조종하는 무인 무기나 무인항공기 등이 여기에 해당한다. 2단계 ‘human on the loop’는 무기체계의 전투 행위 과정을 모니터링하면서 무기체계가 인간의 의도와 다른 방향으로 행동하거나 사격 여부를 결정해야 하는 중요한 순간에만 인간이 개입하는 것으로, 신속한 대응이 필요한 방공무기와 같이

6) 조현석, “인공지능, 자율 무기체계와 미래 전쟁의 변환,” 『21세기 정치학회보』 제28집 제1호(2018), p. 118.

7) US DoD, Directive on Autonomy in Weapon Systems, Nr.3000.09(Nov. 21, 2012)

상당한 수준의 자율성을 보유하는 무기가 여기에 속한다.⁸⁾ 3단계 ‘human out of the loop’는 완전한 자율성을 가진 무기체계로서 인간의 개입이 필요없이 스스로 표적을 찾아 교전 여부를 결정하고 타격하는 무기체계이다. 통상 자율 살상무기(lethal AWS, LAWS) 또는 킬러 로봇으로 불린다. 이 수준의 자율무기가 윤리적으로나 국제법적으로 가장 논란의 대상이 되고 있다. 현재까지는 개발되지 않았으나 인공지능의 발전추세로 보아 머지않아 등장할 것으로 보인다.⁹⁾

앞에서 제시된 자율성의 범주는 사실 자율 무기체계에 내장된 인공지능 기술에 의한 사고(orient)와 판단(decide) 능력을 기준으로 분류한 것으로, 공학적 요소인 로봇 기술(mechanics)에 해당하는 관측(observe)과 행동(act) 능력은 고려되지 않아 자율 무기체계의 포괄적 특징을 설명하기에는 완전하지 않다. 왜냐하면 자율 무기체계 또는 지능형 전투로봇은 인공지능 두뇌와 로봇 몸체로 이루어져 있기 때문이다. 오늘날 인공지능의 중요성이 부각되면서 자율 무기체계에 대해서도 두뇌의 자율성이 강조되는 측면이 있다. 그렇지만 이에 못지않게 중요한 것이 자율적 사고와 판단 결과를 행동으로 옮기는 기능이다. 자율 무기체계에 내장된 인공지능 능력으로 표적을 식별하고 정확한 판단을 하더라도 적절히 대응하지 못하거나 의도하는 대로 작동하지 않는다면 자율 무기체계 또는 지능형 전투로봇이라고 하기에는 부족한 면이 있다. 따라서 자율 무기체계는 지능적 능력에 추가하여 로봇의 특징인 기계적, 공학적 능력이 추가되어야 할 것이다. 대표적인 기계적 능력으로는 이동성(mobility)을 들 수 있다. 원하는 목표지점으로 스스로 이동하고 위험지역이나 장애물을 회피하며 임무 종료 후에는 기지로 스스로 귀환할 수 있어야 한다. 또한 스웜링(swarming)과 같이 다른 무인체나 인간 전투원과 네트워크화하면서 협조된 작전을

8) 미국의 팔랑스(Phalanx) 체계는 실시간으로 자료를 수집, 잠재적 표적을 탐지-평가-추적한다. 함정에 접근하는 물체가 함정을 공격하기 위한 것으로 판명된 경우 긴급 모드에서 갑판에 설치된 캐틀링건이 자동으로 공격을 가할 수 있다. 해상 팔랑스 시스템의 지상 버전인 C-RAM(the Counter Rocket, Artillery and Mortar System)은 敵의 미사일이나 로켓 등의 공격을 탐지, 확인하고 이를 방어하기 위한 무기체계로서 2005년 이라크에서 처음 배치되었다. 이스라엘의 아이언돔(Iron Dome), 독일의 자동 방어시스템인 엔비에스 만티스(Mantis), 네덜란드의 골키퍼(Goalkeeper)나 러시아의 카쉬탄(Kashtan)도 근접방어 무기로서 레이더 표적을 추적 및 탐지한 후 피어싱별을 통해 표적을 스스로 요격할 수 있다. 미국의 패트리엇(Patriot)나 이지스 체계(Aegis Combat System) 또한 적군의 탄도 미사일 공격을 방어하기 위한 것으로 자동모드나 긴급모드로 사용시 자율 무기체계로서 기능을 수행한다. 이스라엘의 하피(Harpy)는 무인기와 순항미사일을 결합한 무기체계로 순찰지역을 자율적으로 운항하다가 敵 레이더를 탐지하며 고성능 폭탄으로 표적을 공격·파괴할 수 있다. 하피보다 길이가 조금 긴 하롭(Harop)은 공격 중지가 능이 탑재되어 있어 인간 운영자는 의도하지 않은 공격을 회피하기 위해 작동을 중지할 수 있다. 예를 들어 저비용 자율공격체계(Low-Cost Autonomous Attack System)는 2003년 자율적으로 목표물을 찾아 공격하는 시험에 성공했으나 인간 운영자의 개입이 없는 전술적 운영의 문제점을 인식하고, 인공지능을 통해 통제가 가능한 human in the loop 방식으로 전환되었다.

9) 여기서 말하는 완전 자율무기체계는 탐지에서 공격까지 인간의 개입 없이 자동으로 이루지는 팔랑스나 패트리엇의 긴급 모드와 같은 수준을 의미하지 않으며, 인간의 선택적 조종이 아닌 전적으로 기계의 자체 판단에 따라 작동되고 어느 정도의 이동성(mobility)과 상호운용성(interoperability)을 지닌 무기체계를 말한다.

수행할 수 있는 상호운용성(inter-operability) 능력이 있어야 한다. 이러한 기계적, 공학적 능력은 고정 발사대나 함정 등에 고정되어 있어 위치를 스스로 변경하지 못하는 미사일 발사체계가 표적 탐지부터 사격에 이르기까지 전 과정을 자율적으로 수행할지라도 이를 지능형 전투로봇이라고 부르기 어색한 이유이기도 하다. 이 글에서 관심을 가지는 자율 무기체계는 바로 이러한 공학적 능력을 지닌 지능형 무기체계를 말한다.

3. 자율 무기체계의 한계와 문제점

인공지능과 로봇의 기술이 발전할수록 자율 무기체계가 고도화되어 결국 인간의 통제를 벗어나 기계가 인간을 지배할 것이라고 우려하는 목소리가 커지고 있다.¹⁰⁾ 특히, 3단계 수준의 완전한 자율 무기체계는 핵무기에 버금갈 정도의 파괴력을 지니게 될 것으로 보인다. 영화 터미네이터와 매트릭스에서는 고도로 지능화된 기계가 인간을 공격하고 인류 문명을 파괴하며, 인류는 이들과 맞서 처절한 전쟁을 벌인다. 이러한 영화와 같은 상황이 반드시 오지는 않을지라도, 어느 시점에 자아를 가진 강인공지능이 개발되어 스스로 판단하고 통제 불가능한 수준에 이르게 된다면 인류 문명에 예측할 수 없는 부정적 영향을 주게 될 것이다.¹¹⁾ 특히 군사용 인공지능과 로봇, 이를 결합한 자율 무기체계는 본질적으로 전쟁의 수단으로서 살상과 파괴를 전제하고 있어 자칫 잘못하면 치명적인 흉기가 되어 아군이나 국민에게 막대한 피해를 줄 수 있다는 점에서 특별한 주의와 대책이 요구될 것이다. 여기서 군사용 인공지능과 로봇, 즉 자율 무기체계의 몇 가지 한계와 잠재적 문제점을 제시한다.

첫째, 인공지능은 입력된 데이터와 알고리즘에 따라 작동하기 때문에 데이터의 질에 큰 영향을 받는다. 빅데이터 전문가 캐시 오닐(Kathy O'Neil)이 그의 책 『Weapons of Math Destruction』에서 지적하듯이 편향된 또는 부정확한 데이터가 입력되면 잘못된 판단과 결정으로 이어지며 잘못된 행동과 치명적 결과를 발생시킨다.¹²⁾ 인간의 편견과 오류, 정보의 부정확성이 그대로 인공지능에 입력되어 결과로 산출된다. 만일 아군에 수집한 적에 대한 데이터가 부정확하거나 불충분할 경우 예상치 못한 결론으로 이어질 수 있다. 적에 대한 정보가 제한적일 수밖에 없으며 가변적이고 불확실한 작전 환경의 본질적 특성을 고려하면 이러한 문제는 계속될 수밖에 없을 것으로 보인다.¹³⁾

둘째, 인공지능에 투입된 데이터와 알고리즘의 신뢰성 문제로서, 블랙박스 속에서 처리되어 제시된 산물을 인간이 어느 수준까지 신뢰할 수 있는가이다. 앞에서 제시된 인공지능 데이터의 부정확성이나 부족, 또는 알고리즘의 편향은 인간이 가지는 본질적인 문제가 인공지능에 전이되는 결과이기도 하다. 다만 인공지능이 인간보다 뛰어나게 유추하고 신속하게 처리할 수 있다는 장점을 가지고 있다. 그러나 인간의 경우 정보처리 과정과 이유에 대한 설명이 가능하여 도출된 결과에 대한 재차 논리적 판단이 가능하지만, 인공지능은 과정에 대한 논리적 설명 없이 결과만 제시하기 때문에 최종적으로 인간이 채택 여부를 결정하는데 어려움이 발생할 것이다. 특히 여러 개의 인공지능 시스템이 통합하여 시스템간 상호작용이 일어날 경우 인공지능이 어느 수준에서 무슨 학습을 했으며, 어떤 이유로 특정의 결론을 도출했는지를 판단하기 더욱 어렵게 만들 것이다.

셋째, 만일 인공지능의 처리 결과를 그대로 인정한다면 이것은 인공지능에게 완전한 자율성을 부여하는 것을 의미하며, 이는 인공지능이나 자율 무기체계는 더 이상 인간의 보조 수단이 아닌 인간과 대등한 차원에서 전쟁 또는 전투 행위의 주체가 될 것이다. 이렇게 되면 인간이 자율 무기체계에 대한 통제권을 상실하여 자율 무기체계가 최초 인간의 의도와는 달리 행동을 하더라도 이를 효과적으로 저지하기 어렵게 된다. 책임 문제에 있어서도 불법적으로 인명을 살상한 로봇이나 기계를 처벌할 수 없고 관련자들에 대한 형사처벌도 거의 할 수 없을 것이다. 인공지능의 기술이 발전할수록 시스템이 더 복잡해지고 인공지능의 역할과 행동 범위가 확장됨에 따라 개발자나 운용자의 의도와 다른 방향으로 행동하거나 그 결과도 예측하기 어려워질 가능성이 커질 것이다. 이에 따라 인공지능(또는 자율 무기체계) 행위의 결과에 대한 책임의 귀속이 모호해져 윤리적 주체가 사라질 가능성이 있다.¹⁴⁾ 인공지능(또는 자율 무기체계) 자체를 윤리적 행위자로 규정하면 개발자나 운용자의 윤리적, 법적 책임은 사라지게 된다. 그렇다고 운전자 또는 지휘관에게 형사적 책임을 묻기도 쉽지 않다. 왜냐하면, 개인이 실제적 또는 추정적으로 자동무기의 오작동을 알고도 시정조치를 제대로 하지 않았거나 고의 또는 과실로 발생한 것이 아니라면 처벌이 어렵기 때문이다.¹⁵⁾

넷째, 인공지능은 기본적으로 데이터와 디지털 네트워크 체계를 기반으로 하는데 이는 적의 사이버 공격이나 전자기적 공격에 치명적일 수 있다. 인공지능이 사이버 공격이나 방어에 매우 효과적인 기술이면서도 이와 동시에 사이버 공격에 취약한 모순을 가지고 있다. 적이 아군의 데이터 시스템이나 클라우드에 침투하여 데이터를 조작하거나 명령체계를 교란하여 아군을 적으로 오인하도록 하거나 인공지능 시스템을 전자기적으로 파괴하거나 전자 바이러스로 작동불능 상태로 만들어 자율 무기체계가 무용지물이 될 수도 있다. 네트워크로 연결된 자율 무기체계가 적의 수중에 들어가 알고리즘이나 네트워크의 취약점이 노출되면 치명적 결과가 발생할 것이다. 단말 자율 무기체계가 많아지면 네트워크 노드가 급증하기 때문에 이러한 가

10) Peter Asaro, "On banning autonomous weapon systems: human rights, automation, and the dehumanization of lethal decision-making", International Review of the Red Cross, vol. 94 Number 886 Summer(2013), p. 688.

11) Ray Kurzweil, op cit.: Amnon H. Eden and James H. Mor. Singularity hypotheses: A Scientific and Philosophical Assessment. (Dordrecht: Springer, 2012).

12) Kathy O'Neil, Weapons of Math Destruction (New York: Crown, 2016)

13) 2016년 마이크로소프트사의 인공지능 챗봇 테이(Tay)가 인간과의 채팅을 통해 비속어와 거짓말, 인종적 편견, 파시스트적 역사관을 학습하여 그대로 따라 하기 시작하여 큰 충격을 주고 운영이 중단되었다. <https://goenair.com/news/article.html?no =2129> (검색일: 2021.7.18)

14) Allen Collin et al. Robot Ethics: the Ethical and Social Implications of Robotics, (Mass: MIT Press, 2011)

15) 박문연, "자율무기체계의 발전과 우리 군의 현실에 대한 고찰," LAW & TECHNOLOGY, 제2권 제6호.(2016) pp. 51-67.

능성은 더욱 커질 것이다. 또한 적의 EMP 공격으로 한순간에 모든 인공지능과 자율 무기체계의 기능이 마비되는 상황도 배제할 수 없다.

다섯째, 기술적 비대칭성에 따른 전쟁의 유혹과 피해의 확산이 더 커질 수 있다는 점이다. 인공지능과 자율 무기체계가 무장한 국가는 그렇지 못한 국가의 취약점을 활용하여 신속하고 효과적으로 공격할 수 있어 전쟁의 유혹에 쉽게 노출될 가능성이 크다.¹⁶⁾ 양측 모두 인공지능으로 무장된 자율 무기체계가 강화된 군사력을 가지고 있다면 군사적 균형이 이루어져 전쟁이 억제될 수 있겠지만, 그렇지 않으면 군사적 균형이 무너지게 된다. 자율 무기체계가 무장한 국가와 그렇지 못한 국가 간의 전쟁은 기계와 인간의 전쟁이 될 것이며, 자율 무기체계를 가진 국가는 자국의 인명 손실에 대한 부담이 없어 전쟁을 너무 쉬운 것으로 만들어 인류의 파괴가 확대될 것이다.

4. 인공지능 기반의 미래전에 대비하기 위한 선행 과제

인공지능과 로봇 기술 기반의 무기체계가 가지는 여러 가지 부정적 효과가 제기되고 있음에도 불구하고 무인화, 자율화의 미래전 양상은 거스를 수 없는 시대적 대세일 것이다. 사실 자율 무기체계가 마냥 부정적인 면만 있는 것은 아니다. 자율 무기체계는 매우 정밀·정확하여 불필요한 살상을 방지하고 인간과 같은 실수를 하지 않아 오히려 인명 피해를 줄일 수 있다. 또한 감정이나 스트레스가 없어 전쟁에서 흔히 발생하는 인간의 비윤리적 행위로 인한 무고한 인명 살상과 파괴를 예방할 수 있을 것이다.¹⁷⁾ 따라서 인공지능 기반의 자율 무기체계가 가지는 부정적 효과를 최소화하면서 인공지능과 로봇 기술을 군사적으로 적절히 활용한다면 피아간의 인명 손실을 최소화하면서 전투를 수행할 수 있는 효과적인 방안이 될 것이다. 이를 위해 대비해야 할 선행 과제로 다음과 같이 몇 가지를 제시하고자 한다.

첫째, 부정확하고 불충분한 데이터로 인해 발생하는 인공지능의 오류를 최소화하기 위한 빅데이터 기반 체계의 구축이다. 인공지능의 능력이 갑자기 길러지는 것이 아니다. 이는 풍부하고 질 좋은 데이터를 활용한 꾸준한 자기 학습을 통해서 달성된다. 따라서 평소부터 인공지능을 학습할 수 있는 분야별로 빅데이터가 축적되어야 하며, 동시에 기계학습을 위한 데이터의 질(정확도)을 검증하는 절차를 마련해야 한다. 더욱 중요한 것은 인공지능 알고리즘 개발자와 데이터 관리자의 융합이 필요하다. 현재는 군사보안으로 인해 인공지능 개발자의 데이터 접근이 어려워 효과적인 알고리즘 개발이 쉽지 않을 뿐 아니라 인공지능의 기계학습도 어려운 실정이다. 기관별 정보 독점과 폐쇄적이고 고압적인 군사보안은 인공지능의 군사적 적용에 가장 큰 걸림돌이 되고 있어, 이에 대한 제도적 해결책이 시급하

다.

둘째, 인공지능의 신뢰성을 높이고 인공지능이 제시하는 결과를 운용자가 채택하기 위해서는 인공지능이 판단한 이유를 설명할 수 있어야 한다. 조건반사적인 신속한 대응이 필요한 경우가 아니라면 인공지능의 정량적 분석에 인간의 논리적 타당성을 추가하여 인공지능의 신뢰성을 향상할 수 있다. 이는 또한 인공지능의 판단 결과와 인간의 상식이나 직관이 불일치할 때 발생할 수 있는 선택의 혼란을 방지하는 방법이기도 하다. 특히 상황분석이나 지휘 결심용 인공지능의 경우에는 설명 가능한 인공지능(Explainable AI, XAI)의 개발이 요구된다. 이는 인공지능의 편향 오류를 수정하는 방법이기도 하다. 최근 초보적인 설명 가능한 인공지능이 개발되고 있다.¹⁸⁾ 이것을 군사 분야에 적용하여 인공지능의 산출물에 인간 운용자의 논리적 이해를 추가한다면 인공지능의 오류를 효과적으로 교정할 수 있을 것이다. 이것은 특히 지휘관의 의사결정을 보조하는 인공지능 참모에게 있어서 반드시 갖추어야 할 기능이다.

셋째, 인공지능의 돌발적 행위나 운용자가 의도하지 않은 행위로 인한 불법적 결과가 발생하지 않도록 하기 위해서는 자율 무기체계가 언제나 인간의 직·간접적인 통제하에 있어야 한다. 완전한 자율 무기체계라 하더라도 인간 전투원과 마찬가지로 독자적인 행동을 하지 않도록 해야 한다. 인간 전투원이 단독으로 활동하지 않고 항상 리더의 지휘하에 팀 단위로 움직이듯이 자율 무기체계도 인간 리더의 지휘하에 팀 단위로 활동을 통해 인간들과 접촉을 유지하며 지속적인 상호작용을 하도록 인간 전투원과 로봇 전투원이 혼합 편성되어 훈련과 작전에 투입되어야 한다. 또한 자율 무기체계가 허용 범위 내에서 자율적 행동을 하도록 허용하되, 인간의 지휘를 벗어나 독자적인 행동을 하면 인간 운용자 또는 지휘자에 의해 즉각적으로 제지될 수 있는 기술적 제재 수단이 강구되어야 한다. 이것은 모자이크 전투 구상에서 제시되고 있는 인간 지휘(command), 기계 통제(control) 개념의 일부이다. 기존의 전투지휘 방식이 인간(지휘관)에 의한 지휘 통제(command & control)로 일원화되어었다 한다면, 미래의 지휘 통제는 인간과 기계가 각각 지휘와 통제를 나누어 행사하는 새로운 유형의 분권화 지휘 기법이 개발되어야 한다.

넷째, 인공지능 기반 무기체계의 태생적 취약점인 사이버 해킹이나 전자기적 공격에 의한 치명적 피해로부터 오는 작전 수행의 차질을 최소화하기 위해서는 모든 의사결정 시스템이나 무기체계를 수동과 자율의 이중(dual) 모드 기능을 갖추어야 한다. 사이버 제공권을 확보한 상태에서는 자율 모드로 운용하다

18) 미국 방위고등연구계획국(DARPA)은 지난 2016년 XAI 투자 프로그램을 발표했으며 국내 UNIST를 포함한 세계의 유명 대학 및 연구 기관 등도 XAI 개발에 박차를 가하고 있다. 미국 심머신(simMachines, Inc)은 AI 심머신(AI simMachines)이라는 XAI를 개발하여 '해석의 제시'라는 기준에 양립이 어렵다고 알려진 AI 활용 과제를 자체 알고리즘으로 해결하였다. 구글 또한 오랫동안 설명 가능한 XAI를 개발해왔고, 클라우드 오픈링을 통해 이 서비스를 출시하기도 했다. IBM은 오픈 스케일(Open Scale)과 함께 설명 가능한 AI 서비스를 제공하고 있다. 이 서비스는 모델이 어떻게 작동하고 있는지를 명확히 할뿐 아니라 관련된 문제를 해결하는 데도 도움을 준다. <http://www.aitimes.kr/news/articleView.html?idxno=14859>(검색일: 2021.7.16)

16) James Acton, "Escalation through Entanglement: How the Vulnerability of Command and Control Systems Raises the Risks of an Inadvertent Nuclear War," *Defense & Security Analysis* Vol. 35 No. 2 (2019).

17) Christopher Toscano, "Friend of Humans: An Argument for Developing Autonomous Weapon Systems", *Journal of National Security Law & Policy* (2015), p. 2.

가 유사시에는 수동 모드로 전환하여 인간이 직접 조작하여 작동될 수 있도록 기술적인 대비책이 강구되어야 하며, 자율시스템의 가동이 중단되는 상황에 대비하여 핵심 기능을 대체할 전문인력을 보유하고 있어야 한다. 이는 무인 자율시스템이 대세인 상황에서도 왜 인간 전투원과 전문인력이 필요한지를 말해 준다. 또한 전략임무를 수행하는 대부대는 물론이고, 창끝 전투를 담당하는 말단 단위 제대들도 사이버와 우주 등의 다영역 작전을 수행할 수 있는 능력을 갖추으로써 말단 제대의 자율 무기체계를 사이버 공격으로부터 보호할 수 있다.

다섯째, 인공지능과 자율 무기체계로 무장한 국가의 전쟁 유흥과 대량 파괴를 방지하기 위해서는 상대적인 비대칭성을 허용해서는 안 된다. 적국 기계의 공격을 인간 전투원이 방어한다면 그 국가는 일방적인 인명 피해를 보고 전투의지를 조기에 상실하게 될 것이다. 처음부터 전쟁을 수행할 수 없는 일방적인 상태가 될 것이다. 이러한 극단적 사태가 발생하지 않도록 하기 위해서는 자율 무기체계의 양적, 질적 균형을 유지해야 할 것이다. 이것이야말로 미래의 전투력이 자율 무기체계를 중심으로 재편되어야 하는 이유이다. 단지 형식적으로 소규모의 자율 무기체계를 갖추는 수준이 아니라 미래전의 게임 체인저 역량을 확보할 수 있는 수준이 되어야 한다. 이것은 앞으로 심각한 병력 부족에 직면할 한국군이 나가야 할 방향이기도 하다.

5. 결론

인공지능과 로봇 기반의 무기체계에 대한 부정적인 효과가 제기되고 반윤리적인 자율 무기체계의 개발을 금지하자는 국제적 여론이 있지만, 미래의 전쟁은 점차 인공지능과 로봇 기술을 적용한 무인화, 자율 무기체계를 중심으로 수행될 것이라는 점은 명확해 보인다. 과거 기관총이 등장하자 기관총의 무자비한 살상력에 놀란 사람들이 전쟁에서 기관총 사용을 금지하고자 하였지만, 기계 사용이 대세인 시대의 흐름을 거스르지 못하고 실패하였듯이 자율 무기체계도 그러할 것이다. 그렇다고 한다면 자율 무기체계의 한계를 극복하고 전장에서 효과적으로 활용할 수 있는 방안을 마련하여 새로운 형태의 미래전에 대비하는 동시에 인명의 피해를 최소화하면서 전쟁을 효과적으로 수행할 수 있어야 한다. 이것은 인구절벽 시대에 한국군이 직면한 병역자원 부족 문제를 해결할 수 있는 길이기도 하다.

자율 무기체계가 인간 전투원을 대신하여 전투를 수행하도록 하기 위해서는 이로 인해 발생할 수 있는 여러 가지 부정적인 효과를 예방할 수 있는 대책을 사전에 마련해야 할 것이다. 첫째, 인공지능에 투입되는 데이터의 부족, 왜곡, 오류를 해소할 수 있어야 한다. 둘째, 인공지능의 판단과 조언에 대한 신뢰성과 채택 가능성을 높이기 위해 인공지능의 산물에 대한 논리적 설명이 가능하게 해야 한다. 셋째, 자율 무기체계의 독단적 활동을 하지 못하도록 인간 리더의 통제하에 팀 단위 활동을 하도록 유무인 복합으로 전투 편성하고, 자율 무기체계의 돌발 행동시 즉각 제지할 수 있는 물리적 장치가 마련되어야 한다. 넷째, 자율 무기체계가 사이버 공격이나 전자기적 공격에 노출되지 않도록 하되, 불가피할 때 대비하여 무기체계의 기능을

이중화하여 필요시 인간이 직접 수동으로 조작할 수 있도록 하며, 핵심기능을 대체할 전문인력을 확보해야 한다. 다섯째, 자율 무기체계는 상대국과의 비대칭적 불균형이 발생하지 않도록 충분히 보유해야 한다. 특히 인구절벽으로 인해 심각한 병력 부족에 직면한 한국군에게 있어 자율 무기체계는 부족한 병력을 상쇄할 수 있는 대안이기도 하다.

References

- [1] Allen Collin et al. Robot Ethics: the Ethical and Social Implications of Robotics, Mass: MIT Press, 2011.
- [2] Amnon H. Eden and James H. Mor. Singularity hypotheses: A Scientific and Philosophical Assessment. Dordrecht: Springer, 2012.
- [3] Christopher P. Toscano, "Friend of Humans: An Argument for Developing Autonomous Weapon Systems", Journal of National Security Law & Policy. Vol. 189, 2015.
- [4] Cho, Hyun Seok, "Artificial Intelligence, Autonomous Weapon System and the Transformation of Future War," 21st Century Political Science Review, Vol. 28, No. 1, pp. 115-139, 2018.
- [5] Kathy O'Neil, Weapons of Math Destruction (New York: Crown, 2016).
- [6] James Acton, "Escalation through Entanglement: How the Vulnerability of Command and Control Systems Raises the Risks of an Inadvertent Nuclear War," Defense & Security Analysis Vol. 35, No. 2, 2019.
- [7] Park, Mun-eon, "Study on Development of Autonomous Weapon System and Reality of ROK Military," LAW & TECHNOLOGY, Vol. 12, No. 6, pp. 51-67, 2016.
- [8] Peter Asaro, "On banning autonomous weapon systems: human rights, automation, and the dehumanization of lethal decision-making", International Review of the Red Cross, vol. 94 Number 886 Summer 2013.
- [9] Ray Kurzweil, Singularity is near, New York: Penguin Books, 2006.
- [10] US DoD, Directive on Autonomy in Weapon Systems, No. 3000.09 (Nov. 21), 2012.
- [11] <https://goeonair.com/news/article.html?no=2129> (검색일: 2021.7.16.)
- [12] <https://100.daum.net/encyclopedia/view/14XXE0074082> (검색일: 2021.7.16.)
- [13] <http://www.aitimes.kr/news/articleView.html?idxno=14859>(검색일: 2021.7.16.)