Vincenzo Pisani

Difesa, sicurezza, incertezza

L'impiego dell'Intelligenza Artificiale nel warfare

L'ingresso dirompente di questa tecnologia nel panorama globale della difesa è destinato a determinare la gestione dei futuri conflitti insieme agli equilibri di potere e alle alleanze fra i diversi Stati. Siamo a un punto di svolta. Denso di insidie. La sfida è già in corso su palcoscenici rilevanti e riguarda specifiche applicazioni militari. Come, altresì, forme di manipolazione e falsificazione di foto e video. Tuttavia, vi è un aspetto più inquietante e destabilizzante. L'attuale impossibilità delle potenze militari a conoscere l'effettiva consistenza dei sistemi difensivi e offensivi sul terreno supportati dall'IA. Ma allora, come impostare un nuovo percorso di deterrenza come successe nella fase della Guerra Fredda con la corsa all'armamento nucleare? Oggi, al tempo dell'asimmetria esistente tra sistemi autocratici e democratici, questa è la domanda delle domande.

Gli analisti militari d'oltreoceano definiscono le tecnologie dirompenti nel settore della difesa come "insieme di tecnologie applicate a un problema rilevante in un modo che altera radicalmente la simmetria del potere militare tra i concorrenti" e che "valica immediatamente le politiche, le dottrine e l'organizzazione di tutti gli attori coinvolti".
Il loro impiego nel panorama globale della difesa sta generando numerosi e sostanziali quesiti riguardo la gestione dei futuri conflitti e, più in generale, degli equilibri di potere fra Stati. Fra tutte, una in particolare rappresenta il vero punto di svolta, lo spartiacque tra passato e futuro.

Nonostante non esista una definizione univoca e condivisa di Intelligenza Artificiale, generalmente il termine IA viene utilizzato per riferirsi a un sistema informatico con capacità cognitive a livello umano. L'IA è divisa in due categorie: IA ristretta e IA generale. Rientrano nella prima categoria quei sistemi che possono eseguire solo il compito specifico per il quale sono stati addestrati. I secondi, tramite apprendimento autonomo, potrebbero un giorno essere in grado di eseguire una vasta gamma di compiti, compresi quelli per i quali non sono stati specificamente addestrati.

L'IA ristretta è attualmente integrata in numerose applicazioni militari, che includono – ma non sono limitate – a intelligence, sorveglianza e ricognizione, logistica, operazioni informatiche, comando e controllo e sistemi semi-autonomi e autonomi.

La vulnerabilità a distorsioni cognitive e bias

I sistemi abilitati dall'IA potrebbero reagire molto più velocemente di quelli che si basano sull'input dell'operatore, far fronte a un aumento esponenziale della quantità di dati disponibili per l'analisi e abilitare nuovi concetti operativi. Ne è un esempio il risultato evidenziato nel programma di ricerca del DARPA "AlphaDogfight", incentrato sulle capacità di combattimento aereo tramite

IA. In una serie di duelli aerei simulati fra un velivolo pilotato tramite IA e un altro operato da un pilota umano, il primo ha prevalso con risultati sorprendenti.

Tuttavia, l'IA apre anche una serie di sfide trasversali. Prima fra tutte, è la vulnerabilità a distorsioni cognitive e bias. Diversi ricercatori hanno ripetutamente individuato casi di pregiudizi razziali nei programmi di riconoscimento facciale, principalmente dovuti alla mancanza di diversità nelle immagini su cui i sistemi sono stati allenati. Allo stesso tempo, alcuni programmi di elaborazione del linguaggio hanno sviluppato pregiudizi di genere.

Vulnerabilità che, in ambito militare, potrebbero portare a conseguenze letali. Ad esempio, incorporare inconsciamente pregiudizi non rilevati in fase di test potrebbe condurre a casi di identificazione errata dei bersagli.

A ciò va aggiunto il ruolo dell'IA nel consentire falsificazioni e manipolazioni digitali di foto, audio e video con risultati sempre più realistici. Queste capacità dell'IA possono essere impiegate – e sta già avvenendo – come operazioni mirate a minare le capacità informative del nemico, per influenzare le opinioni pubbliche e creare instabilità nei sistemi sociopolitici di Paesi ritenuti avversari.

Per questo motivo, alcuni analisti sostengono che le piattaforme di social media, oltre a impiegare strumenti di rilevamento dei deep fakes, dovrebbero rafforzare le soluzioni di classificazione e autenticazione dei contenuti.

L'errore da evitare: una fiducia inappropriata nel sistema

C'è poi un altro tema critico, che riguarda l'"esplicabilità": un concetto che integra le idee di "intelligibilità" e "responsabilità" e allude alla trasparenza nella progettazione e nel processo interno che gli algoritmi seguono per la selezione e l'elaborazione dei dati, a partire dai quali i sistemi stabiliscono modelli e prendono decisioni. Le tipologie di algoritmi di IA che hanno le prestazioni più elevate non sono attualmente in grado di spiegare i loro processi. Specie in un contesto militare, poiché l'opacità nel funzionamento dell'algoritmo potrebbe indurre gli operatori ad avere una eccessiva o scarsa fiducia nel sistema e mettere in discussione diversi passaggi dell'interazione uomo-macchina.

Il primo è senza dubbio l'allineamento degli obiettivi: l'uomo e la macchina devono avere una comprensione comune dell'obiettivo. In un ambiente dinamico, gli obiettivi cambiano e sia l'uomo che la macchina devono adattarsi simultaneamente sulla base di un quadro condiviso dell'ambiente corrente. Il secondo passaggio fondamentale è l'allineamento dei compiti: l'uomo e la macchina devono capire i confini dello spazio decisionale dell'altro, specialmente quando gli obiettivi cambiano. In questo processo, gli esseri umani devono essere perfettamente consapevoli dei limiti del progetto della macchina per evitare di riporre una fiducia inappropriata nel sistema. Infine, non meno importante, è il passaggio che riguarda l'interfaccia uomo-macchina: a causa del requisito di decisioni tempestive in molte applicazioni militari dell'IA, le interfacce tradizionali possono rallentare le prestazioni. È quindi necessario considerare delle soluzioni per garantire un coordinamento in tempo reale fra uomo e macchina.

Infine, l'impiego dell'IA nel warfare pone una questione ancor più dirimente. Se ogni progresso nello sviluppo della tecnologia bellica ha portato con sé l'incertezza sui suoi possibili impieghi e sulla sua potenza, l'IA introduce un grado di incertezza ancora maggiore. All'inizio della Guerra Fredda, sia gli Stati Uniti che l'Unione Sovietica erano a conoscenza delle capacità distruttive delle armi nucleari e temevano che l'avversario potesse svilupparne di più potenti. Il risultato fu la corsa agli armamenti nucleari.

L'Intelligenza Artificiale, dal canto suo, genera entrambe le forme di incertezza: nessuno sa ancora esattamente come le armi abilitate dall'IA saranno usate sul campo di battaglia, tanto

meno quanto saranno potenti. Piuttosto che costituire un singolo sistema d'arma in sé, l'IA può essere incorporata in molteplici sistemi e infrastrutture come i centri di comando e controllo e nelle soluzioni logistiche. Eppure, non è facile determinare come queste innovazioni cambieranno la natura dei conflitti bellici. Che effetto avranno sciami di sottomarini senza equipaggio sulla guerra navale? Cosa succederà quando l'IA non sarà solo integrata negli armamenti e nei centri di comando e controllo esistenti, ma inserita in essi tramite un processo bottom-up? Quale esercito sarà in grado di integrare l'IA in maniera più efficiente e veloce nei suoi sistemi d'arma e nelle sue tattiche e con quale vantaggio sul campo di battaglia? Questa incertezza su come l'IA sarà impiegata e quanto sarà efficace genera, quindi, sfide significanti per gli Stati, in particolare nella strategia militare.

Le tecnologie emergenti e disruptive stanno sfidando il modo in cui la deterrenza, la difesa e, più in generale, le strategie di sicurezza sono formulate e applicate a livello nazionale e multilaterale. Le dimensioni territoriali non rappresentano più il fattore principale per determinare il potere di uno Stato. Lo sviluppo tecnologico, l'agilità di manovra nonché la velocità e l'accuratezza del processo decisionale conteranno più delle risorse a disposizione. Pertanto, un attore di piccole dimensioni – ma con notevoli capacità tecnologiche – potrebbe essere in grado di sfidare con successo una grande potenza.

L'obsoleto paradigma della stabilità internazionale

La situazione è resa ancora più complicata dall'ambiente multipolare della competizione e del conflitto. Un'azione intesa a scoraggiare un avversario potrebbe produrre preoccupazioni inaspettate tra gli altri attori. La deterrenza nucleare ha prodotto stabilità strategica adoperando una combinazione di negoziazione, dichiarazioni pubbliche e programmi mirati all'acquisizione di armamenti.

Nell'ambiente attuale e del prossimo futuro, acquisire più armi non produrrà maggiore stabilità e la capacità di negoziare su questioni strategiche e di controllo degli armamenti con gli avversari è significativamente ridotta rispetto al passato. Trovare un modo per coordinare questo nuovo ambiente strategico e rafforzare la stabilità internazionale non è intuitivo. Infatti, considerando che gli Stati stanno cercando di ampliare la deterrenza contro rischi emergenti e contro nuove armi non nucleari, il vecchio paradigma della stabilità è ormai compromesso e dovrà essere rinnovato tenendo conto delle armi abilitate da tecnologie dirompenti e dei loro effetti sulla deterrenza.

Una piattaforma di dialogo permanente

In particolare, l'IA introduce la capacità di influenzare la deterrenza militare e la coercizione in modi unici. Può alterare i calcoli costi-benefici eliminando la *fog of war*, vale a dire la difficoltà di ottenere informazioni attendibili in situazioni di guerra, imponendo la razionalità sulle decisioni politiche e diminuendo il costo umano dell'impegno militare. Può ricalibrare l'equilibrio tra misure offensive e difensive, facendo pendere l'ago della bilancia a favore della prevenzione, e minare i presupposti esistenti nella deterrenza convenzionale e nucleare. In altre parole, l'IA potrebbe fornire agli utilizzatori la capacità di agire sulla base di informazioni raccolte, sintetizzate ed elaborate in tempo reale, aumentando la certezza e la severità delle strategie di coercizione e comprimendo la distanza tra l'intelligence, le decisioni politiche e l'azione coercitiva.

In linea generale, il contesto del prossimo futuro richiede, quindi, il mantenimento di un primato tecnologico credibile, in grado di alimentare una deterrenza efficace che induca ipotizzabili aggressori a effettuare – prima delle rispettive iniziative – una valutazione costo-beneficio. Senza questo primato, l'asimmetria esistente tra sistemi autocratici e democratici sarà difficilmente mitigata da qualsivoglia diplomazia e forma di diritto internazionale.

Senza un impianto di consapevolezza sarà difficile mantenere una soglia di deterrenza credibile a difesa del sistema valoriale cui le differenti comunità del globo sentono di appartenere, specie se incardinate su una struttura democratica e pacifica.

Il percorso per raggiungere tale consapevolezza non può che passare per una piattaforma di dialogo permanente, che coinvolga tutti gli attori chiave che contribuiscono alla costruzione degli scenari di difesa e sicurezza – governi, forze armate, mondo accademico e della ricerca, imprese – in un quadro di confronto multilaterale. È infatti su tale consapevolezza condivisa che potremo basare l'asse morale e il perimetro di principi etici entro i quali operare e delimitare l'impiego delle nuove tecnologie, e in particolare dell'IA, nei futuri scenari di difesa e sicurezza.

L'articolo prende spunto dal paper "Winning the Artificial Intelligence Era. Quantum Diplomacy and the Power of Automation", realizzato da Fondazione Leonardo-Civiltà delle Macchine in collaborazione con Centro Studi Americani:

https://www.civiltadellemacchine.it/it/la-fondazione/umanesimo-digitale/winning-the-artificial-intelligence-era-quantum-diplomacy-and-the-power-of-automation

Vincenzo Pisani è coordinatore dei progetti di ricerca presso Fondazione Leonardo-Civiltà delle Macchine.