

למידה עמוקה, רובוטים ומערכות נשק

גדעון עקביה¹ וערן יהודה²

האם רובוטים יחליפו את חייל צה"ל בשדה הקרב בעתיד הקרוב? כותבי המאמר עונים על כך בשלילה ומציעים נקודת מבט ביקורתית וספקנית ביחס לאפשרויות המגולמות בשילוב טכנולוגיות מבוססות "אינטליגנציה מלאכותית" בצה"ל. הכותבים מציגים סקירה קצרה של היישומים של טכנולוגיות מבוססות "אינטליגנציה מלאכותית" בתחומים אזרחיים וצבאיים שונים וניתוח של המשמעות הנובעת מכך. זאת, בדגש על הסיכונים הרבים והקשיים הנמשכים בתרגום הפוטנציאל המגולם בטכנולוגיות אלה ליתרונות צבאיים ממשיים בזמן לחימה. לטענתם, בעוד שפיתוח והטמעת טכנולוגיות "אינטליגנציה מלאכותית" בצה"ל אכן יוכל לשפר באופן משמעותי את תפקודי העורף הצבאי (ממחקר מודיעיני ועד לרפואה צבאית), טכנולוגיות אלה לא יוכלו להוות תחליף לכוחות הלוחמים.

"אי אפשר להכיר את הבלתי צפוי"
הרקליטוס, הפרגמנט השמונה עשר

מבוא

"אינטליגנציה מלאכותית" היא התכונה המאפשרת למערכת מעשה ידי אדם להתנהג בצורה המקרינה את מה שהיינו מכנים "אינטליגנציה" אצל בני אדם. מאמר זה דן בשני תתי-תחומים של "אינטליגנציה מלאכותית" (AI): למידה עמוקה (Deep Learning) המבוססת על נתונים רבים (Big Data) מצד אחד, ורובוטים ומכוניות אוטונומיות (המסתמכים בין השאר על למידה עמוקה), מצד שני. תחומים אלו מושכים בשנים האחרונות תשומת לב רבה בין החוקרים ואנשי העסקים. גם בצה"ל ובמערכת הביטחון בכלל, מושקעת מחשבה רבה באפשרויות לשלב יישומים של AI בתוכניות לבניין הכוח.

¹ ד"ר גדעון עקביה, פיזיקאי ואיש מחשבים העוסק בחקר הקשר שבין חדשנות טכנולוגית לבין בניין כוח צבאי והפעלתו.

² סגן-אלוף (מיל') ערן יהודה, שירת עד לאחרונה כחוקר ביצועים במינהל למחקר, פיתוח אמצעי לחימה (אמל"ח) ותשתית טכנולוגית ולפני כן כראש תחום חקר ביצועים במחלקת אמל"ח של זרוע היבשה.

מאמר זה מבקש לבחון את היכולת של מערכות ה-AI לסייע לצה"ל לנצח במלחמה העתידית, תוך שמירה על חייהם של לוחמינו. נציג סקירה ביקורתית של מערכות הלמידה העמוקה השונות הקיימות, ובאופן ספציפי נבחן את יעילותם של רובוטים וכלים אוטונומיים, המסתמכים על יכולות מסוג זה, עבור שדה הקרב העתידי. אל מול בחינת ההזדמנויות, נבקש להציף את הסיכונים הרבים, את הפערים בטכנולוגיות הקיימות והצפויות, ואת הקשיים בשילוב מערכות AI ב"ממלכת אי-הוודאות", שהיא המלחמה. קשה ליישם מערכות AI שתפעלנה ביעילות נוכח אויב פעיל, המיישם יוזמות מפתיעות. בעוד שמערכות AI מסייעות כבר בימינו לחוקרים בגופי מודיעין עורפיים ולאנשי מפקדה העוסקים ב"קרב המפה", אנו נסביר מדוע מערכות AI תתקשנה מאוד לעזור ללוחמים בקצה המבצעי, העוסקים בניהול "קרב המגע". קל יחסית לעזור לחוליית תצפית על ידי חיישנים חכמים, ולעזור לצלף על ידי כוונת חכמה, אבל קשה מאוד להפוך רובוט ללוחם רב-תכליתי, המסוגל לבצע מגוון גדול של משימות.³

אם להרוס את סוף הסיפור, רובוטים אכן יופעלו בקרב, בעתיד הנראה לעין, אך אנו טוענים כי למרות הטכנולוגיה המתפתחת והאופק המבטיח, רובוטים יופעלו בקרב רק בתפקידים מוגדרים ומצומצמים. **רובוטים "קרביים"** לא יחליפו, כמובן, את הלוחמים בסמטאות השטח הבנוי, הם **לא יהיו אוטונומיים, וידרשו מפעיל אנושי שיוקדש למשימה**, מפעיל שבמהלך לחימה קרקעית יידרש ברוב המקרים להימצא בקרבת הרובוט. את שילוב הרובוטים בלחימה צריך יהיה לתכנן בצה"ל בהתאם למגבלות התקציב וכוח האדם.

מודלים מתמטיים פשוטים ו-AI

הרעיון שאפשר להשתמש בכלים מתמטיים כדי לשפר את יכולתנו להסיק מסקנות מנתונים העומדים לרשותנו, אינו חדש. הוא הועלה כבר לפני למעלה ממאתיים שנה, כאשר לג'נדר וגאוס הציעו, כל אחד בנפרד, להשתמש בשיטת "הריבועים המינימליים", *least square*, בכדי לנבא, על סמך תצפיות חלקיות, את מסלולם הצפוי של גופים אסטרונומיים. היום זמינים ברשת כלי תוכנה חינוכיים המאפשרים לבנות בקלות מודלים מתמטיים "אינטליגנטיים" מסוגים שונים.

³ על האבחנה בין "קרב המפה" ל"קרב המגע" ראה עקביה, גדעון. "שו"ב ממוכן בכוחות השדה – בעיות יסוד". **מערכות** 407, יוני 2006. קרב המפה הוא תחום הפעילות של אנשים, "כמו קציני המטה במפקדת אוגדה, המרכזים את רוב תשומת ליבם במפה. המפה מייצגת בעבורם את השדה הקרב, ועליה הם מסמנים את המידע המגיע אליהם"; קרב המגע הוא תחום הפעילות של "הלוחמים המעורבים בלחימה ואשר מפעילים נשק בכיוון ישיר מתמקדים בעיקר בסביבתם הישירה". אנחנו נפנה תשומת לב בעיקר לתרומה האפשרות של מערכות AI לקרב המגע, בניגוד לרוב המאמרים שהתפרסמו לאחרונה בישראל. ראו למשל נורית כהן-אינגר וליאור זכרון, " 'לראות את הנולד' – שילוב יכולות בינה מלאכותית לשיפור העוצמה האווירית של צה"ל", **בין הקטבים** 11-12, 2016. כהן-אינגר, נורית וגל א' קמינקא, "והרי התחזית: צה"ל בדרך לצבא תבוני – מפת דרכים לאימוץ טכנולוגיות בינה מלאכותית בצה"ל", **בין הקטבים** 18, דצמבר 2018. דוד שורש, "אסטרטגיה בעידן הבינה המלאכותית: תהליך תכנון אסטרטגי למגמות טכנולוגיות משבשות", **בין הקטבים** 18, דצמבר 2018.

אבל גם כאשר ישנם נתונים זמינים שמהם אפשר לכאורה להתחיל, לא מובטח שהנתונים הזמינים רלוונטיים, וגם מערכות בינה מלאכותית נוטות לעיתים, כמו אנשים, להיתפס לנתונים לא חשובים. לעיתים, בהיעדר מידע רלוונטי, מסתפקים מפתחי המערכות במידע חלופי ובנתונים "ממלאי מקום". בארה"ב, למשל, קיימים מזה שנים גופים מסחריים האוספים מידע על כל אדם, מחשבים בעזרת מידע זה את "ציון האשראי" שלו, ומוכרים אותו לכל דורש. לא ניכנס כאן לדיון בשאלה אם "ציון האשראי" המחושב אמנם מנבא בצורה טובה את הסיכוי שאדם יחזיר הלוואה, אבל מכיוון שציון האשראי זמין בקלות למי שמוכן לשלם בעבורו, מקובל להשתמש בו כנתון "ממלא מקום" לאפיון אנשים למטרות אחרות, גם אם התוצאה מעוותת. בפלורידה, למשל, אדם עם ציון אשראי נמוך והיסטוריה טובה של נהיגה ישלם בעבור ביטוח רכב \$1500 לשנה יותר מאשר אדם עם ציון אשראי גבוה והרשעות על נהיגה במצב שכרות. כלומר, בפלורידה מסתפקים במידע חלופי ובנתונים "ממלאי מקום" גם כאשר יש מידע רלוונטי יותר.⁴

לפעמים קל לפתח מודל שיסייע בקבלת החלטות, אבל קשה לבדוק את איכותו. אם המודל משרת גוף מסחרי, טבעי להניח שהוא ייבחן, ושהמודל ומפתחיו יקבלו משווא חוזר המאפשר למודל "ללמוד" ולשפר את ביצועיו. חברת "אמזון", למשל, בוחנת ומעדכנת כל הזמן את המודלים שלפיהם היא מציעה ללקוחות מוצרים שהם עשויים לרצות לקנות.

גופים ציבוריים ולא מסחריים עושים לעיתים שימוש במודל כדי להבטיח בסיס נוח לקבלת החלטות שיהיה קשה למתוח עליהן ביקורת, ולא כדי להבטיח את איכות ההחלטות. במקרה כזה לא חייבים לעדכן את המודל ולספק לו משווא. במילים אחרות, לא חייבים להבטיח שהמודל טוב - מספיק שהוא מקובל וקל להפעיל אותו. לפעמים גם כאשר יש מידע המספק משווא על ביצועי המודל, לא משתמשים בו כדי לא לחשוף את המערכת לביקורת.⁵

מערכות "למידה עמוקה"

מערכות הלמידה העמוקה - Deep Learning - הנמצאות בשימוש רב בעת האחרונה הן עתירות נתונים, כלומר מתבססות על מה שמכונה Big Data. בעזרת נתונים רבים ניתן להגיע לתוצאות טובות אפילו ללא למידה עמוקה. כפי שהסבירו חוקרים אנשי גוגל, במאמר שפורסם כבר בשנת 2009, גם מודלים פשוטים, שאינם מנסים "להבין" את המציאות ושאינם מנסים לממש מודעות מלאכותית, יכולים, אם הם מתבססים על נתונים רבים מאוד, לתת תוצאות טובות, טובות באופן "בלתי סביר".⁶

⁴ C. O'Neil. *Weapons of math destruction: how big data increases inequality and threatens democracy*. Broadway (2016). pp.137-138.

⁵ Ibid, pp.98-100.

⁶ Halevy, Alon, Norvig, Peter, and Pereira Fernando. "The Unreasonable Effectiveness of Data". *IEEE Intelligent Systems* 24, (Mar-Apr 2009). no. 2, pp. 8-12.

אין זה מקרה, כמובן, שמאמר זה נכתב בידי אנשי גוגל. הרעיון המקורי של מפתחי מנוע החיפוש היה לשלב חיפוש של מילות מפתח עם ההנחה שדפים פופולריים יתנו תשובה "טובה" או רלוונטית יותר לשאלות של המשתמשים. את פופולריות הדפים השונים ברשת הם העריכו לאחר שוטטות ברשת וספירת מספר ההפניות לכל דף.⁷ השימוש הנרחב במנוע החיפוש מזהה בעבור גוגל את הדפים שנמצאו רלוונטיים, וכל מי שמשתמש שוב ושוב במנוע החיפוש מספק לגוגל מידע על הנושאים המעניינים אותו כאדם. עוצמתה האימתנית של גוגל מתבססת על השילוב בין יכולתה להציע לציבור המשתמשים שירותי חיפוש טובים לבין יכולתה להציע למפרסמים שירותי פרסום ממוקדים, המותאמים לכל אדם בהתאם לתחומי העניין שלו.

תוכנת התרגום של גוגל התבססה גם היא בתחילה על קיומן של דוגמאות לתרגומים טובים ברשת, דוגמאות המאפשרות לתוכנה לזהות דרך מקובלת לתרגום משפטים גם בלי להבין אותם. האיחוד האירופי, למשל, מפיץ את מסמכיו הרשמיים בשפות רבות במקביל, וכך גם ברית נאט"ו. כבר בשנת 1990, כאשר הציעה קבוצה של חוקרים מ־IBM לבסס תרגום מכונה על כלים סטטיסטיים, הם השתמשו ברשומות הפרלמנט הקנדי, שהעמידו לרשותם מאה מיליון דוגמאות המופיעות באנגלית ובצרפתית גם יחד. בשנת 2014, כאשר התכוננו לתחרות בין תוכנות תרגום, כבר התבססו על מאות ואלפי מיליונים של דוגמאות.⁸ מנוע החיפוש של גוגל ותוכנת התרגום שלה שופרו ושוכללו כמובן מאז הופיעו לראשונה, בין השאר בעזרת מה שמכונה רשתות נוירונים ולמידה עמוקה.

כבר בשנות החמישים של המאה העשרים היה מי שהציע לבנות רכיב חומרה דמוי נוירון – כלומר שיהיו לו שני מצבים, פעיל או לא פעיל – והסביר שרשת כזו של נוירונים תוכל לבצע משימות כמו זיהוי עצמים בתמונה, מבלי שהגורמים המשפיעים יוגדרו בעבורה מראש ומבלי שמתכנת אנושי יחליט איך לפעול בכל הנסיבות האפשריות.⁹ אבל החומרה הזמינה בשנות החמישים לא הייתה חזקה מספיק למימוש מועיל של הרעיון, ורוב החוקרים נטשו אותו. חלק מהאטרקטיביות המקורית של הרעיון נבעה מהשימוש במונח נוירון, ומהתקווה כי רשת נוירונים תוכל לשחזר תפקודים של המוח האנושי. אבל הנוירונים במוח האנושי הם ישויות מורכבות בהרבה, והם בעלי יותר משני מצבים.

בשנות השמונים פותחו אלגוריתמים לעדכון המבנה של רשתות נוירונים בהתאם

⁷ Lawrence, Page, Brin, Sergey, Motwani, Rajeev and Winograd, Terry. "The Page Rank Citation Ranking: Bringing Order to the Web"; Brin, Sergey and Lawrence, Page (1998). "The anatomy of a large-scale hypertextual Web search engine". *Computer Networks and ISDN Systems* 30, (29 Jan. 1998), pp. 107-117.

⁸ Peter F. Brown, John Cocke, Stephen A. Della Pietra, Vincent J. Della Pietra, Fredrick Jelnek, John D. Lafferty, Robert L. Mercer, and Paul S. Roossin. "A Statistical Approach to Machine Translation, *Computational Linguistics* 16, no. 2, pp. 79-85 (June 1990); Ian Goodfellow, Yoshua Bengio and Aaron Courville. *Deep Learning*. MIT Press, (2016) p.19.

⁹ Frank Rosenblatt. "The perceptron: A probabilistic model for information storage and organization in the brain". *Psychological Review* 65, (1958), pp. 386-408.

לתפקודם המבוקש. הרעיון הבסיסי היה שאם מאמנים רשת נוירונים על ידי מתן דוגמאות רבות של הבעיה שהיא מתבקשת לפתור, בצירוף התשובה הנכונה, כלומר בעזרת אספקת "נתונים מתויגים", אז הרשת תוכל להתאים את עצמה, ללא מגע יד אדם, כדי לבצע טוב יותר את משימתה.¹⁰ הצעד הקריטי הבא בדרך ליישום הנרחב של למידה עמוקה התבסס על התובנה שהמעבדים החזקים המיוחדים, המכונים GPUs - Graphical Processing Units, שפותחו למשימות של עיבוד גרפי, מתאימים מאוד גם לביצוע החישובים הנדרשים לאימון רשת נוירונים. בעזרת ה-GPUs אפשר לממש רשתות נוירונים גדולות ועמוקות, כאלה המכילות שכבות רבות של נוירונים, אשר תיישמה "למידה עמוקה".

בניגוד למודלים הפשוטים שאותם הזכרנו קודם, התלויים בכך שהגורמים המשפיעים יזוהו מראש בידי מפתחי המודלים, יכולות מערכות הלמידה העמוקה לזהות בעצמן את הגורמים המשפיעים ולהחליט בעצמן איך להשתמש בהם. האדם מעורב באספקת הדוגמאות המתויגות, בהגדרת התוצאות הרצויות ובקביעת נקודות ההתחלה לפעילות האלגוריתם. אבל הפרמטרים הנבחרים ברמות הביניים, הקובעים למעשה את המאפיינים המזוהים ואת הגורמים המשפיעים, אינם נבחרים על ידי אדם אלא נלמדים ישירות במערכת, מהדוגמאות שסופקו לה. מערכות לפענוח תמונות, למשל, יכולות לחסוך זמן לרופאים ולסייע להם. כדי לאפשר זאת מועסקים, בעולם השלישי, מאות אלפי אנשים, שאינם רופאים, בתיוג התמונות הנדרשות כדי לאמן מערכות אלו.¹¹

שלושה מהחוקרים הבולטים בתחום של למידה עמוקה, שקיבלו לאחרונה את פרס Turing על תרומתם לנושא, הסבירו בשנת 2015 כי במערכת למידה עמוקה טיפוסית, מאפיינים את רשת הנוירונים מאות מיליונים של פרמטרים, ולכן נדרשות מאות מיליונים של דוגמאות (מתויגות!) לאימון רשת הנוירונים ולכילולה.¹² מערכות כאלו יכולות לבצע בהצלחה משימות מוגדרות היטב, למשל, לזהות את העצם המופיע בתמונה, ולהבחין בין תמונות רבות שאינן כוללות מידע חדש לבין התמונות המעטות הכוללות מידע חדש ומעניין. מערכות למידה עמוקה יכולות לשחק בהצלחה משחק, אפילו מורכב, כאשר חוקי המשחק מוגדרים וברורים, ולזהות מילות מפתח בתוך דיבור.

¹⁰ David E. Rumelhart, Geoffrey E. Hinton and Ronald J. Williams. "Learning representations by back-propagating errors", *Nature* 323, (9 Oct. 1986), pp. 533-536.

¹¹ Cade Metz. "A.I. Is Learning from Humans. Many Humans: Artificial intelligence is being taught by thousands of office workers around the world. It is not exactly futuristic work" (16 Aug. 2019). *New York Times*; Economist (4 Jan. 2020). "Technology in China," *The Economist Technology Quarterly*, pp. 9-10.

¹² Yann LeCun, Yoshua Bengio and Geoffrey Hinton. "Deep learning: Review". *Nature* 521, (28 May 2015), pp. 436-444.

אם בתחילה התבסס מנוע החיפוש של גוגל רק על מילות המפתח בשאלתה, הרי עכשיו הוא מיישם למידה עמוקה כדי לנסות להבין את השאלתה כמשפט. זאת ועוד, גוגל כבר איננה מסתפקת בהצבעה על דפים פופולריים ברשת, ומעסיקה כשישה עשר אלף בוחני איכות, הפועלים בהתאם למסמך הנחיות מפורט, שמהדורתו מדצמבר 2019 כוללת 168 עמודים.¹³ כאשר תוצאות של חיפוש מסוים אינן באיכות הנדרשת הדבר מצביע לעיתים על צורך בשינוי אלגוריתם החיפוש. האלגוריתם המעודכן מתחשב, למשל, במקומו של השואל, מעדיף לתת הפניות לדפים "טריים", וגם מנסה, לאחרונה, לתת עדיפות לדפי מידע שמקורם ידוע כמהימן.

גם אלגוריתם התרגום של גוגל (Google Translate) השתפר מאוד מאז יציאתו לשוק. אם בתחילה הוא היה מבוסס רק על השכיחות של חלקי משפט שונים, אזי היום הוא מנסה לנתח משפטים שלמים, בעזרת למידה עמוקה. אבל מהעובדה שאלגוריתם התרגום של גוגל פועל במקרים רבים בצורה טובה, לא נובע שמתאים לייחס לו יכולת הבנה. הסטטיסטיקה המגולמת ברשת העמוקה, שעליה מבוססת תוכנת התרגום, משקפת רק קשר בין מילים למילים, ואין בה שם רמז לקשר בין מילים לרעיונות או למושגים מופשטים. התוכנה של גוגל יכולה לתרגם נכון כמעט את כל המילים, אבל לפעמים תפספס לחלוטין את המסר של המשפט.¹⁴

למערכות למידה עמוקה יש לעיתים ביצועים מרשימים, אך הן סובלות מכמה נקודות תורפה. הן עלולות לשקף הטיות באוסף הדוגמאות שעליו הן מתבססות. גם כאשר הן מצליחות לזהות הן מתקשות להבין, ואפשר להטעות אותן על ידי שינויים קלים אך ממוקדים בקלט שלהם, או על ידי נתונים או צעדים מסוג בלתי צפוי.

מכיוון שמערכות למידה עמוקה מתבססות על כמות גדולה מאוד של דוגמאות מתויגות, קשה לבצע בקרת איכות של הדוגמאות, בעיקר אם הן נאספות עצמאית מהרשת. מכיוון שמבנה רשתות הנוירונים העמוקות והגדולות שמיישמות את מערכות הלמידה העמוקה איננו שקוף למשתמש, לא קל לזהות הטיות כאלה. **אם משתמשים במערכת באופן עיוור והיא בעלת הטיות שיטתיות, אזי הדבר עלול להוביל לתוצאות חמורות.**

דוגמה להטייה שיטתית קיימת, למשל, במאגרי התמונות המשטרתיים בארה"ב שכללו לאחרונה כ-117 מיליון תמונות. רוב הגופים העוסקים בשמירת החוק אינם בודקים את דיוק המערכות לזיהוי פנים שבהן הם משתמשים, ואינם מפעילים בקרה אנושית על תוצאות ההשוואה, אם כי ברור שבקרה כזאת נדרשת. המשטרה נוטה לעצור יותר שחורים, כך שטעויות בזיהוי ישפיעו בעיקר עליהם. לכן יש אנשים,

¹³ Google. "Search Quality Evaluator" (5 Dec. 2019). <https://guidelines.raterhub.com/searchqualityevaluatorguidelines.pdf>

¹⁴ Drew McDermott. "Artificial Intelligence Meets Natural Stupidity", *SIGART* (28 May 2015); Douglas R. Hofstadter. *Le ton beau de Marot: in praise of the music of language* (1997): Basic; Lewis-Kraus Gideon, "The Great A.I. Awakening: How Google used artificial intelligence to transform Google Translate", *New York Times* (14 Dec. 2016); Douglas Hofstadter, "The Shallowness of Google Translate". *The Atlantic* (30 Jan. 2018).

כמו נשיא מיקרוסופט, הקוראים לרגולציה ממשלתית של בסיסי מידע ותוכנות המשמשים לזיהוי פנים.¹⁵

גם מערכות המצליחות לזהות עצמים בודדים מתקשות לעיתים לטפל בתמונה הכוללת עצמים רבים. המשימה של "הבנת" תכולתה של תמונה ותיאור תכולה זו באופן מילולי היא משימה קשה הרבה יותר מ"זיהוי תמונה", כלומר משיוכה לאחת הקטגוריות בקבוצה של קטגוריות המוגדרות מראש. יש הטוענים כי עדיף לא לנסות לתאר את התמונה בשלבים – קודם זיהוי ואחר כך בניית תיאור מילולי – אלא להשתמש במערכת למידה עמוקה שתתרגם ישירות את התמונה המוצגת לתיאור המילולי המתאים. למערכת לתרגום ישיר כזה שהם פיתחו יש הצלחות רבות, אבל יש לה גם כישלונות. למשל, תמונה אחת זוהתה נכון, כקבוצה של צעירים המשחקים ב"צלחת מעופפת", ואילו האחרת זוהתה בטעות כמקרר המלא במזון.¹⁶ הכנסת שינויים קלים וממוקדים לתמונות יכולה לגרום למערכות למידה עמוקה לטעות בזיהוי, על אף שלאנשים אין קושי לזהות אותן נכונה גם אחרי השינויים הללו.

רוב התוצאות המדגמיות את רגישותן של מערכות AI לשינויים קלים אך ממוקדים, התקבלו במעבדות של החוקרים ולא בסביבה תפעולית מציאותית. על-כן הן לא ממצות את הנושא אך מצביעות על כך שיש מקום לדאגה. לעיתים אפשר להטעות מערכת למידה עמוקה גם באמצעים פשוטים. למשל, מערכת גילוי אשר גילתה אדם אחד בצד אחד של התמונה אך לא גילתה את האדם בצד השני, בגלל "טלאי" שהוא נושא על גופו.¹⁷

דוגמה אחרת לאמצעי הטעיה פשוט נתן לאחרונה אמן גרמני שהטעה את תוכנת הניווט של גוגל על ידי הולכת עגלת צעצוע שבתוכה הניח 99 טלפונים סלולריים. התוכנה הניחה שהיא רואה 99 כלי רכב שבחרו לנסוע דווקא בנתיב מסוים, הכוונה את הנהגים האחרים בהתאם למסקנות מ"עובדה" זו, ויצרה פקק תנועה מיותר.¹⁸ אנו רגילים לכך שעוצמת החישוב של הטלפון הסלולרי שלנו מספיקה לכל מה שאנחנו עושים בחיי היום-יום, אבל יישומן של מערכות למידה עמוקה דורש יכולת חישוב חזקה וכמויות אנרגיה רבות הרבה יותר, והדבר יגביל, כמובן, את היכולת

¹⁵ Joy Buolamwini and Gebru Timnit. "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification". *Proceedings of Machine Learning Research* 81, (2018), pp. 1-15; Georgetown Law. "The Perpetual Line-Up: Unregulated Police Face Recognition in America". Center on Privacy and Technology (18 Oct. 2018); Brad Smith (Microsoft president). "Facial recognition technology: The need for public regulation and corporate responsibility", *Microsoft Blog* (6 Dec. 2018); Wilson Benjamin, Judy Hoffman, and Jamie Morgenstern. "Predictive Inequity in Object Detection" (21 Feb. 2019). (*arXiv:1902.11097v1 [cs.CV]*).

¹⁶ Oriol Vinyals, Alexander Toshev, Samy Bengio and Dumitru Erhan. "Show and Tell: A Neural Image Caption Generator" (20 Apr. 2015). (*arXiv:1411.4555v2 [cs.CV]*).

¹⁷ S. Thys, W. Van Ranst and T. Goedemé. "Fooling automated surveillance cameras: adversarial patches to attack person detection", (2019), *arXiv:1904.08653*.

¹⁸ Brian Barrett. "An Artist Used 99 Phones to Fake a Google Maps Traffic Jam", *Wired*, (3 Feb. 2020).

ליישם למידה עמוקה בשטח. לענקי האינטרנט כמו גוגל או פייסבוק יש מרכזי חישוב ענקיים, הממוקמים במקומות שבהם יש אנרגיה זולה, ויש טענות כי הדבר עלול להטות את כיווני המחקר ולמנוע מעורבות של שחקנים קטנים, כמו חברות הזנק או אוניברסיטאות.

יש אנשים המקווים לראות בעתיד מערכות AI חזקות הרבה יותר מאלו שאנו מכירים היום – מערכות שתדענה להתמודד עם מציאות עמומה, בלתי מוכרת ומורכבת, לבנות מודלים סיבתיים של המציאות, ולהסביר בצורה משכנעת איך הגיעו למסקנות שלהן. אבל הדעה המקובלת היא שאנחנו עדיין רחוקים מהיכולת לבנות מערכות כאלו. מערכות הלמידה העמוקה המוצלחות המוכרות היום מניחות בדרך כלל שמאפייני העיקריים של העולם אינם משתנים, ודורשות תחזוקה רבה כאשר מתברר שהעולם דווקא משתנה. בעוד שאנשים נהנים, במקרים רבים, מהבנה אינטואיטיבית של העולם החומרי ושל התנהגותם של אנשים אחרים, למערכות למידה עמוקה אין הבנה אינטואיטיבית כזו של המציאות.

שימושים רפואיים במערכות למידה עמוקה

מאמצים רבים הושקעו ועדיין מושקעים במערכות AI רפואיות, וחלקם אף הוכתרו בהצלחה. מערכות למידה עמוקה לניתוח תמונות ולפענוח יכולות להגדיל את הסיכוי לזיהוי מוקדם של מחלות שונות ולעיתים לסווג טוב יותר מרופא אנושי.¹⁹ בית חולים אחד התחיל לאחרונה לבחון את השימוש במערכת למידה עמוקה כדי לזהות במהירות חולי קורונה כבר בחדר המיון על סמך תצלום הריאות שלהם. בעוד שקבלת תשובה מרנטגנולוג מומחה שיפענח את התצלום דורשת לרוב שש שעות, הרי קבלת החלטה ממערכת הלמידה העמוקה היא כמעט מיידית.²⁰

תיאור השימושים הרפואיים במערכות AI תורם לדיון שלנו כי הוא מצביע על הקושי באיסוף ובתיגון נתונים רלבנטיים, ועל קשיים תפעוליים אחרים המפריעים ליישום מערכות AI. ההבדל בין מערכות היכולות לשרת בהצלחה את הרופאים במרכזי המחקר הגדולים לבין מערכות היכולות לשרת בהצלחה את האחיות במרפאות הכפריות, דומה מבחינות רבות להבדל בין מערכות המופעלות בגופי המחקר הביטחוניים העורפיים לבין מערכות שצריכות לשרת את החיילים המעורבים בקרב.

יש המקווים שמערכות עיבוד תמונות רפואיות תוכלנה לעזור להתגבר על המחסור ברופאים באזורים כפריים. גוגל פיתחה, למשל, מערכת למידה עמוקה שהפגינה בניסויים יכולת לזהות תוך עשר דקות סיבוך של סכרת העלול לגרום לעיוורון, ולעשות זאת בדיוק של 90%, שזו רמת הדיוק של מומחה אנושי. אבל

¹⁹ Philipp Tschandl et al. "Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study". *The Lancet: Oncology* 20, no. 7, (1 July 2019). pp. 938-947.

²⁰ Karen Hao. "Doctors are using AI to triage covid-19 patients. The tools may be here to stay". *Technology Review*, (23 Apr. 2020).

כאשר המערכת נבחנה בניסוי נרחב במרפאות כפריות בתאילנד התברר שהמצב טוב פחות. בניסוי בשטח התמונות היו באיכות ירודה בהרבה מאיכות התמונות בהן השתמשו החוקרים, ובמקרים רבים המערכת סירבה לתת תשובה, ודרשה להפנות את הנבדק לרופא. גם איכות תשתיות האינטרנט במרפאה הכפרית לא הספיקה, לעיתים קרובות, כדי לתקשר עם המערכת המרוחקת, ובזבז זמן רב על ביצוע בדיקות שלא נתנו תשובה. עוד כדאי להזכיר שלא מתאים לבחון את המערכת רק על סמך הדיוק שלה. גם רופאים אנושיים מגיעים, לעיתים קרובות, למסקנות שונות. מערכת מועילה צריכה להשתלב עם מסקנותיה בתהליך שבו אנשי הצוות הרפואי דנים בתוצאות, בלי לקבל אותן כתורה מסיני.²¹

האם אפשר להכניס מערכות למידה עמוקה לשימוש נרחב ברפואה? לכאורה קיימים נתונים רבים, אבל במידע המוזן יש הרבה טעויות רפואיות ופערים, ורוב הרשומות הממוחשבות הזמינות, בארה"ב לפחות, מיועדות לצורך הנהלת החשבונות וחיוב החולים, או חברות הביטוח, ולא לצורך קבלת החלטות רפואיות.²² וגם כאשר יש נתונים רבים, צריך לבדוק האם הם קשורים לעניין, בלתי תלויים, ומייצגים את האוכלוסייה.

מערכות "למידה עמוקה" שאינן מנסות להבין את סיבת הדברים ומסתפקות בניתוח הנתונים עלולות לעיתים להיתפס לגורמים לא מהותיים וללמוד להחליט על-פי גורמים שאינם קשורים למצבו של החולה. מערכות כאלו עלולות, בסופו של דבר, לבסס את החלטתן על נתונים שוליים כמו סוג מכונת ה-MRI שהיה בשימוש, השעה שבה בוצעה בדיקת הדם, או האם החולה זכה לביקורו של כומר. דוגמה מעניינת לכך ניתן לראות בבית החולים "הר סיני" בניו יורק, שם הופעלה מערכת ייעודית לזיהוי, בעזרת צילומי רנטגן, של חולים העלולים לחלות בדלקת ריאות. בעוד המערכת עמדה במשימתה ב"הר סיני", היא נכשלה כאשר בחנה תמונות מבתי חולים אחרים. התברר שהמערכת פשוט למדה ב"הר סיני" להבחין בין חולים שצולמו במחלקת הרנטגן לבין חולים שצולמו במכשיר נייד שהובא אל מיטתם. החולים שלדעת הרופאים לא היה רצוי לטלטלם אל מחלקת הרנטגן נטו להיות החולים הקשים, ומערכת הלמידה העמוקה קישרה עובדה זאת לשימוש במכשיר הרנטגן הנייד. אבל ההתבססות על השימוש במכשיר נייד איננה מייצגת כמובן תובנות מהתצלום על מצבו של החולה. אם בבית חולים אחר נוהלי השימוש במכשירים ניידים הם אחרים, אז כמובן שהשימוש במערכת הזו לא יוביל לזיהוי מוצלח של

²¹ Emma Beede. "Healthcare AI systems that put people at the center" (25 Apr. 2020). *Blog. google*; Emma Beede, Elizabeth Elliott Baylor, Fred Hersch, Anna Iurchenko, Lauren Wilcox, Paisan Ruamviboonsuk and Laura M. Vardoulakis. "A Human-Centered Evaluation of a Deep Learning System Deployed in Clinics for the Detection of Diabetic Retinopathy" (Apr. 2020). Paper presented at the CHI: Conference on Human Factors in Computing Systems.

²² Liz Szabo. "Artificial Intelligence Is Rushing into Patient Care - And Could Raise Risks". *Scientific American*, (24 Dec. 2019).

החולים הקשים שם.²³ לאור העובדה שקשה להעביר מערכת למידה עמוקה אפילו מבית חולים אחד למשנהו, יש להתייחס בספקנות לתקווה שנוכל לפתח מערכות למידה עמוקה אשר תפעלנה בהצלחה במלחמה עתידית, כאשר המציאות איננה יציבה וכל שדה קרב שונה מקודמו.

שימושים צבאיים במערכות למידה עמוקה עתירות נתונים

אם הדברים שהבאנו כאן אמנם משקפים את מצב הטכנולוגיה של למידה עמוקה, מה הלקח שצריך להפיק מכך? מהם השימושים הצבאיים במערכות למידה עמוקה עתירות נתונים שבהם כדאי לנו להתמקד בישראל? לצורך הדיון בשאלה במה כדאי להתמקד נבחין בין שימושים בשגרה לבין שימושים בלחימה, ובה נבחין בין שימושים בנוהל הקרב לבין שימושים בניהול הקרב.

מערכות למידה עמוקה בשגרה

טבעי לשלב מערכות למידה עמוקה ב"קרב המידע", כלומר בפעילותם של גופים, כמו גופי המודיעין המרכזיים והעורפיים, העוסקים בשגרה ולאורך זמן באיסוף מידע רב ובניתוחו. בתהליך ההפקה של המודיעין יש משימות רבות, כמו ניתוח תצלומי אוויר, תמלול הקלטות שמע או תרגום מסמכים, שלביצוען אפשר להיעזר במערכת למידה עמוקה. במאמצים אלו ממילא מעורבים אנשים רבים ולכן קל יחסית לתייג את הדוגמאות, אם כי גודלו של בסיס המידע עלול להיות קטן ממה שמקובל היום בלמידה עמוקה. אפשר להיעזר בלמידה עמוקה גם כדי לקבוע סדר עדיפויות לתמלול ולתרגום, וכדי לנסות ולזהות את הידיעות הנראות מעניינות יותר. אם נוכל לעזור לאנשי המודיעין בביצוע ממוכן של המשימות הללו הם יוכלו להתרכז במשימות העיליות יותר, כמו ניתוח הידיעות החשובות והבנת משמעותן. למידה עמוקה יכולה לשרת, וכבר משרתת, גם את מי שעוסק בלוחמת סייבר, שם צריך להתמודד במהירות עם כמות גדולה של מידע משתנה, או את מי שעוסק בניהול הספקטרום האלקטרומגנטי. כמו כן, אפשר לנצל מערכות AI בכלל ומערכות למידה עמוקה בפרט לשיפור וייעול ההדרכה, אבל אנו לא נעסוק בכך במאמר זה.

מערכות למידה עמוקה בלחימה

בנוהל הקרב: קציני המטה היושבים ליד שולחן במפקדת האוגדה או החטיבה, ואשר עוסקים ב"קרב המפה", שמחים להפוך את המחשב ואת מערכות התוכנה החכמות לכלי עיקרי שלהם, כלי המציג לא רק את המפות אלא גם את תוצרי המודיעין הרבים הזמינים היום. קצינים אלו, יוכלו כאמור לקבל עזרה מעוזרי מחשב חכמים אך מוגבלים, המיועדים, למשל, לניתוח שטח, כדי להצביע על נתיב התקדמות נוח

Ibid. ²³

או כדי להעריך את פרק הזמן שיידרש לביצוע תנועה מנקודה א' לנקודה ב' בנתיב נתון. עזרים ממוחשבים אלו אינם דורשים למידה עמוקה, כי הם מנצלים בסיסי מידע מוגבלים, פיסיקליים בעיקר, שאותם אפשר להכין מראש.

קשה הרבה יותר לעזור למפקדות, על ידי מערכות למידה עמוקה, בביצוע משימות עיליות יותר, כמו הערכת דרכי הפעולה של האויב או בחירת דרך פעולה מועדפת לכוחותינו, מכיוון שבמקרים אלו לא יעמוד לרשותנו המידע המתויג הרב הנדרש לאימון מערכת למידה עמוקה. כאמור, בגופי המודיעין המרכזיים, ברמה המטכ"לית ובפיקוד המרחבי, טבעי לאסוף בשגרה מידע רב על האויב ולהניח שהאויב יתקשה להסתיר לאורך זמן את שגרת החיים שלו. לעומת זאת, אין סיכוי שנוכל לבנות אוסף גדול של דוגמאות מתויגות המראות מה התוצאות של תכנון מבצעי אפשרי כזה או אחר, דוגמת תכנון הלחימה בגזרה הלבנונית בכלל או תכנון כיבוש הבופור בפרט. ביישומים אזרחיים אפשר לבחון דרכי פעולה שונות בעזרת סימולציות ממוחשבות, שאותן קל יחסית להריץ פעמים רבות. אך זה אפשרי רק אם ניתן להעריך מראש את התפלגות המשתנים המשפיעים, ואינו ישים נוכח אויב פעיל ויוזם, העשוי להעמיד אותנו נוכח הפתעות מבניות.

בניהול הקרב: תוך כדי הלחימה עלינו להתמודד עם אויב פעיל, מגיב ויוזם. כדי להבין בזמן אמת מה האויב עושה עכשיו ומה כוונותיו לעתיד, אי אפשר להסתפק במידע על מה שהאויב עשה בשגרה. בתנאי לחימה האויב יתאמץ להפתיע ולהונות אותנו, והוא עשוי לבחור בדרכי פעולה חדשות ולא צפויות ולאמץ שיטות פעולה שלא נחשפו קודם. היישומים הפופולריים של למידה עמוקה מצליחים לסווג מידע רב לקטגוריות ברורות, המוגדרות בדרך כלל מראש. אבל בשדה הקרב הבעיה החשובה עלולה להיות זיהוי דפוסי פעילות חדשים ובלתי מוכרים והבנת מאפייניהם העיקריים. לכן, ההצלחות הגדולות של מערכות למידה עמוקה אזרחיות אינן מוכיחות שמערכות למידה עמוקה תוכלנה להתמודד עם אויב אנושי, פעיל ויוזם. במיוחד קשה לאסוף ולעבד, תוך כדי לחימה, את המידע הרך והחלקי על כוח העמידה של האויב ועל נחישותו, ועל יכולתם של כוחותינו להמשיך להילחם, מידע העשוי להיות קריטי לניהול הקרב. גם כאשר יש מידע זמין, צריך לבחון את מהימנות הנתונים ואת סבירותם.

רוב היישומים האזרחיים הפופולריים של למידה עמוקה אינם ערוכים היטב כדי להתמודד עם "אויבים", ולכן כל משתמשי האינטרנט חשופים לסכנות כמו גניבת זהות או קבלת מידע כוזב. כאשר הבנקים בונים יישומים פיננסיים הם נערכים כמובן להתמודד עם גנבים, אבל הבנקים פועלים בתנאים מוגדרים ומוגבלים, ויסרבו לפעול אם הם יזהו שפעולת המערכת מופרעת או משובשת. תנאי הסביבה שבהם צריך לפעול בשדה הקרב קשים לעיתים הרבה יותר, ואי אפשר להתבסס על מערכת שתחליט להפסיק לפעול בדיוק כאשר התנאים יהיו קשים – כלומר כאשר הלחימה תתחיל להיות רצינית.

אם מתמקדים ברמות הטקטיות הנמוכות, העוסקות ב"קרב המגע", אפשר

לכאורה לאסוף מידע על מספר גדול של אירועים. אבל גם אז קשה להפיק תועלת ממערכות למידה עמוקה, וזאת משתי סיבות:

- א. מגבלות קשב ותשומת לב – בעיצומו של "קרב מגע" מתמקדים הלוחמים בפלוגה בשטח ובהבנת מה שקורה מסביבם. הם יתקשו להתנתק מהשטח ולעסוק בתיוג נתונים שנאספו. אפילו אם נעביר את הנתונים אחורה, אל מישור שיהיה פנוי לתייג אותם, ונרצה להחזיר ללוחמים המעורבים בקרב רק את התוצרים של מערכות הלמידה העמוקה, אנו עלולים לסבול ממגבלות הקשב ותשומת הלב של הלוחמים, שעומס מידע עלול רק לבלבל אותם. אין עדיין "משקפיים" טובים מספיק היכולים גם להציג היטב את הסביבה הקרובה כמו שהיא וגם להוסיף לתצוגה מידע שבא ממערכות מחשב חיצוניות.
- ב. מגבלות משקל, אנרגיה ותקשורת – בדרך כלל אי-אפשר להסתפק בידע שהושג מראש כדי לקבל החלטות בשדה הקרב. אם נרצה לבצע למידה עמוקה במהלך ניהול "קרב המגע" צריך יהיה להעמיד לרשות היחידות הלוחמות, גם ברמות הנמוכות, מערכות תקשוב רבות עוצמה, שנוטות להיות כבדות, ולצורך כמות גדולה של אנרגיה. נוסף על כך, הפעלת מערכות למידה עמוקה דורשת במקרים רבים העברת כמות גדולה של נתונים, ואין עדיין מענה טוב לבעיה של תשתיות תקשורת רחבות-סרט המתאימות לרמה הטקטית הנמוכה, במיוחד לא כאשר נלחמים רגלית.

מכוניות אוטונומיות ורובוטים

אנשי חזון שטענו כי מכוניות אוטונומיות תגרומנה לשינוי מהפכני – במקום לרכוש מכוניות פרטיות גדולות, כבדות, מזהמות ובלתי יעילות, נוכל להתבסס על צי מוניות, שתפעלנה ללא נהג ועל כן תהיינה קטנות, קלות, חשמליות, זולות ופשוטות. בחזון הזה מסתתרים גם אתגרים טכנולוגיים קשים וגם הזדמנויות עסקיות גדולות. לא מפתיע לכן שרבים מיצרני הרכב המסורתיים, וגם גופים כמו גוגל, נכנסו לנושא. אבל אנחנו עדיין רחוקים מהיכולת להפעיל רכבים אוטונומיים לגמרי בכבישים הרגילים, והדבר מודגם מפעם לפעם בתאונות דרמטיות של מכוניות אלו.

דווקא המומחים לדבר מסבירים כי מערכות ה-AI המוכרות לנו היום יכולות לתת מענה טוב רק לבעיות מוגדרות וצרות מאוד. נהיגה, לעומת זאת, איננה בעיה מוגדרת וצרה, היא אוסף גדול של בעיות שאותן צריך לפתור בו בזמן. הרכבים האוטונומיים מסתמכים על למידה עמוקה כדי לזהות את מאפייני הסביבה שבה הם פועלים. הם צריכים לא רק לראות מה קורה מסביבם אלא גם להבין את התנהגותם של הולכי רגל ושל נהגים אחרים – ממשאות ועד אופניים וסקייטבורד – ולנבא מה הם עומדים לעשות. הרכב האוטונומי צריך, למשל, לראות, לשמוע, ולהבין את התנהגותו הצפויה של אמבולנס צופר. אך המכוניות ה"אוטונומיות" הנוכחיות אינן "מבינות" מה קורה בכביש. בנוסף הן מתקשות, למשל, לפעול בתנאי תאורה ומזג אוויר גרועים, או להתמודד עם שינויים מהירים ברמת התאורה.

רוב היצרנים מתקיימים בינתיים במכוניות שלהם כלי עזר לנהג משני סוגים. הסוג הראשון כולל מערכות המתרעות על מצבים העלולים להיות מסוכנים, כמו התקרבות לרכב אחר או סטייה מהמסלול. כלי עזר כאלה אינם מחליפים את הנהג, אבל מפנים את תשומת ליבו לאירועים חריגים ומקטינים בכך את הסיכוי לתאונות. הסוג השני הוא מערכות עזר המשחררות, בתנאים מסוימים, את הנהג מהשליטה הפרטנית בנהיגה, אבל דורשות שהנהג יתפוס פיקוד ברגע שיתעורר הצורך בכך. הדרישה שהנהג האנושי יתפוס פיקוד בדיוק ברגע הנכון היא דרישה קשה, למעשה בלתי אפשרית. הניסיון מראה, שוב ושוב, שהמשתמשים, גם אם הם מומחים לדבר, נוטים להניח שהמערכת האוטונומית פועלת בסדר, ולכן אפשר לסמוך עליה ולהפנות את תשומת הלב לדברים אחרים. המשתמשים ממשיכים להניח שהכול בסדר גם כאשר מתפתח מצב מסוכן, ולעיתים הדבר מוביל לתאונה. זמן התגובה של המפעילים, כאשר הם נדרשים לתפוס פיקוד, הוא ארוך, וכדי להגיב נכון הם זקוקים למודעות מצבית מלאה, שאינה ברשותם אם הפנו קודם לכן את תשומת לבם למשימות אחרות.



תמונה 1: מכונית אוטונומית אשר הייתה מעורבת בתאונה שגבתה חיי אדם (NTSB - National Transportation Safety Board)²⁴

ביוני 2018 התנגש, למשל, רכב "אוטונומי" של טסלה במכשול הפרדה. בחקירת התאונה התברר שהמכונית "התבלבלה" בזיהוי מסלולי הנסיעה ולכן התנגשה במכשול ההפרדה. הנהג, שהיה בדרכו למקום עבודתו, כבר ידע שהמכונית נוטה לטעות בדיוק בנקודה זו בכביש, כי נתקל בכך פעמים אחדות, והוא סיפר על כך קודם. למרות זאת, ולמרות הנחיות היצרן שלפיהן הידיים של הנהג צריכות להישאר על ההגה והוא צריך לפקח על התנהגות המכונית, הוא הפנה את תשומת ליבו ואת ידיו למשחק מחשב.²⁵

למרות שמכוניות אוטונומיות אינן מתאימות עדיין, ולא תהיינה מתאימות בעתיד הנראה לעין, לשימוש בטוח ורב תכליתי של הציבור לצורך נסיעות בכבישים העירוניים הרגילים, יש כבר שימוש במכוניות אוטונומיות בתנאים מוגבלים ומבוקרים. למשל, בהובלת נוסעים מטרמינל אחד בשדה תעופה לטרמינל שני או בתוך מחסני אמזון. דוגמה אחרת לרכבים אוטונומיים המתאימים לשימושים מיוחדים הם הרובוטים הנכנסים לשימוש בחקלאות ומאפשרים לחסוך בכוח אדם, אבל רק אחרי שמכניס את השטח לעיבוד על ידי רובוטים, ומודדים את מיקומו המדויק של כל שיח עגבניות, למשל, בעזרת GPS. בכל מקרה, גם אם רובוט חקלאי כזה יפחית את מספר הפועלים הפשוטים הוא יפעל תחת פיקוח צמוד של אדם המומחה לדבר, שינווט אותו או שילך לידו, וסביר שהרובוטים יגיעו לשדה עם הטכנאים של היצרן.²⁶

מכוונות אוטונומיות ורובוטיות בשדה הקרב

הדיון שלנו במכוניות אוטונומיות וברובוטים חקלאיים קשור כמובן לסוגייה מורכבת הנבחנת בימים אלו – האם אפשר לעזור ללוחמים בקרב המגע בעזרת מכוונות חכמות שהפעלתן לא תדרוש תשומת לב רבה מהלוחמים. כדי לנתח סוגייה זו יש להבחין בין מכוונות המופעלות מרחוק לבין מכוונות אוטונומיות לחלוטין.

מכונה המופעלת מרחוק יכולה לשחרר את הלוחם מהצורך להיכנס לאזור הפעילות, העלול להיות מסוכן במיוחד או מזוהם במיוחד. המשטרה מפעילה, למשל, רובוט לטיפול בעצמים חשודים ולסילוק פצצות. את הרובוט הזה יכולה המשטרה להפעיל ממרחק קטן ובקשר עין, וזה מפשט את הפעלתו ומגביל את תשתיות התקשורת הנדרשות. אבל הפעלת רובוט ב"קרב המגע" מטווה קצר תחייב את המפעיל להתחשב באילוצים הכלליים והקשים המאפיינים את "קרב המגע". ב"קרב המגע" טבעי כמובן להעדיף מכונה המופעלת מטווה רחוק. במקרה כזה

²⁵ Timothy B. Lee, "New Data Backs Up Details in a Fatal 2018 Tesla Model X Crash" (18 Feb. 2020); *Wired*, National Transportation Safety Board, "Tesla Crash Investigation Yields 9 NTSB Safety Recommendations" (25 Feb. 2020). *News Release*; John Thornhill, "Trusting AI too much can be fatal". *Financial Times*, (3 Mar. 2020).

²⁶ John Seabrook, "The Age of Robot Farmers" (15 Apr. 2015). *The New Yorker*; Bloch, Sam. "Strawberry farmers: Come with Harvest CROO if you want to live" (22 Jan. 2019). *The Counter*; Knvul Sheikh. "A Growing Presence on the Farm: Robots". *New York Times*.

צריך להעביר מהמכונה אל המפעיל הרבה יותר מידע, ולשם כך נדרש בדרך כלל לפחות ערוץ וידיאו אחד. כאמור, בשדה הקרב קשה להבטיח את יכולת התקשורת הנדרשת לשם כך, במיוחד אם רוצים להפעיל מרחוק הרבה מכונות שונות באותו זמן. אין היום, וככל הנראה לא יהיה בעתיד הקרוב, פתרון היכול להעמיד לרשות לוחמים ב"קרב המגע" ערוצי וידיאו רבים באותו זמן, אלא אם מסתפקים בתקשורת לטווח קצר.

ואולי אפשר להחליף את הלוחם המעורב ב"קרב המגע" במכונה חכמה אוטונומית לחלוטין? הקשיים במימוש החזון של מכונות אוטונומיות, מימוש המתרחק והולך, מחייבים זהירות בנושא זה. מבחינות מסוימות קל יותר להפעיל רכב יבשתי אוטונומי בשדה הקרב מאשר בעיר, כי בשדה הקרב אין צורך בדרך כלל להתחשב בהולכי רגל, וכי לכלי הרכב הקרביים יש יכולת עבירות טובה יותר מאשר לכלי רכב אזרחיים רגילים. מבחינות רבות אחרות קשה יותר להפעיל רכב יבשתי אוטונומי בשדה הקרב מאשר בכבישים העירוניים בזמן שלום, כי הסביבה מורכבת יותר, במיוחד כאשר נלחמים בשטח בנוי, אז חלק מהתשתיות והמבנים שבעיר עלולים להיהרס וקשה לייצר מפה עדכנית של אזור הלחימה. בנוסף, הלחימה מתנהלת בנוכחות אויב, שיתאמץ למצוא את נקודות התורפה של המערכות האוטונומיות ולשבשן.

חלק מהרעיונות האמריקאיים לשימוש ברובוטים ברמה הטקטית מתמקדים במשימות מוגדרות, לעיתים פשוטות - כמו רובוט שיהיה סָבֵל חי"ר - ולעיתים מורכבות יותר, כמו פריצת מכשול. רובוטים כאלו לא יהיו אוטונומיים, הם ידרשו מפעיל המוקדש לדבר, שלא יהיה מרוחק מאוד מהם, אבל שלא יצטרך להיחשף כמו המפעיל של דחפור אנושי.



תמונה 2: חייל אמריקאי שולט ברובוט בלתי מאויש (צילום: US Army)²⁷

²⁷ https://www.army.mil/article/222153/army_futures_leveraging_mission_command_for_effective_soldier_robot_teams

חלק מההצעות להכניס רובוטים לשימוש נרחב ברמה הטקטית מתעלמות מהעומס שיוטל על הלוחמים המפעילים ומיכולתם המוגבלת לטפל ברובוטים בשדה הקרב. במאמר המתאר חזון לכיתת חי"ר איכותי עתידית, שתהיה עתירת רובוטים מסוגים שונים, שתוכל לנהל קרב מגע בכל סוגי השטח ושגם תהיה לה ניידות אופרטיבית משופרת – מציע המחבר לבנות את הכיתה כך: שנים עשר לוחמים, ארבעה כלי רכב נושאי חי"ר, שמונה רובוטים קרקעיים בעלי ארבע רגליים, שנים עשר רחפנים, כלי רכב אחד לתחזוקת הרובוטים וכלי רכב אחד לניהול סיוע עקיף באש. על אף שכותרת המאמר כוללת את המילים "אמנות האפשרי" חסר בו פירוט למספר האנשים שיידרשו, נוסף על הלוחמים, כדי להפעיל ולתחזק את הרובוטים וכלי הרכב. המאמר גם לא דן בשאלה איך אפשר יהיה להנהיג כיתה מורכבת כזו בשדה הקרב.²⁸

בסך הכול, תוכניות לשילוב של מכונות אוטונומיות ורובוטים עצמאיים בפלוגת החי"ר המעורבת בקרב המגע צריכות להיות תוכניות לטווח זמן רחוק. בעשור או שניים הקרובים, לא נראה רובוטים ומכונות אוטונומיות שיחליפו את הלוחמים בקרב המגע וינהלו בעצמם את הלחימה בסמטאות השטח הבנוי. הרובוטים שייכנסו לשימוש ברמה הטקטית הנמוכה לא יהיו אוטונומיים ויופעלו תחת פיקוח הדוק של מפעיל המוקדש לדבר.

סיכום

אנחנו רחוקים מהיכולת לבנות מערכות אינטליגנטיות מלאכותיות שתדענה לזהות ולהבין את המאפיינים העיקריים של מציאות מורכבת, נוכח אי ודאות ורסיסי מידע סותרים, ושתדענה להסביר בצורה משכנעת איך הגיעו למסקנות ולהמלצות שלהן. המערכות המוצלחות של למידה עמוקה המוכרות לנו היום מתרכזות בביצוע משימות מוגדרות היטב כמו זיהוי תמונה או זיהוי דיבור, ופיתוח דורש כמות גדולה מאוד של נתונים שיתויגו קודם לכן בידי אדם הכשיר לנושא.

טבעי לשלב מערכות למידה עמוקה בגופי איסוף ומחקר מודיעיני, הפועלים בשגרה ולאורך זמן. זאת מכיוון שבשגרה, ובפעילות הביטחון השוטף, אנחנו יכולים לאסוף נתונים רבים, ועשויים להשתכנע שאנחנו יודעים איך האויב פועל ויפעל. עם זאת, בתנאי לחימה עלינו להתמודד עם אויב פעיל, העלול להפתיע אותנו ולאמץ שיטות פעולה שלא נחשפו קודם. מסוכן להניח שיהיו לנו מראש נתונים מפורטים על איך האויב מתכוון לפעול בימים הראשונים של המלחמה הבאה, וגם אם נוכל לאסוף נתונים במהלך הימים הראשונים הללו יהיה קשה להעביר אותם אל מי שיוכל לעבד ולהזין אותם למערכות הלמידה. אפילו אם יהיו לנו דוגמאות רבות מהעבר ומידע רב על ההווה, ואפילו אם מערכות מודיעין המבוססות על Big Data תוכלנה לגלות שקורה משהו חריג, הן תתקשנה להבין מה בדיוק קורה ולזהות את מאפייניו.

²⁸ Jeff Becker. "How Lethal, Mobile, Protected and Aware? Exploring the Art of the Possible in Future Infantry Combat". *Small War Journal*, (8 Nov. 2019).

מערכות למידה עמוקה יכולות בקלות יחסית לעזור לקצינים היושבים ליד שולחן במפקדת אוגדה או חטיבה, אבל מי שינסה להכניס לשימוש מערכות למידה עמוקה אשר תשרתנה את הלוחמים המעורבים ישירות בקרב המגע, יצטרך להתמודד עם מגבלות קשות של קשב ושל תשומת לב. יש רגעים שבהם מפקד פלוגה, למשל, מתכוון למשימתו הבאה ואז הוא ישמח לקבל מודיעין עילי ולקבל עזרה ממערכת שתעזור לו לנתח את השטח ושתצביע בעבורו על נקודות מעניינות בשטח. אפשר לעזור לצלפים ולחוליות תצפית בעזרת מערכות "חכמות", אבל בעיצומו של ניהול הקרב בשטח יתקשו רוב הלוחמים והמפקדים בפלוגה, המשקיעים את כל כולם בהבנת מה שקורה בסביבתם הקרובה, להתנתק מהשטח ולהתרכז בצג מחשב. כמו־כן, מערכות המיועדות לחייל הרובאי צריכות גם להתמודד עם מגבלות קשות של משקל, של אנרגיה, ושל תקשורת.

אם בעתיד הנראה לעין נכניס לשימוש בשדה הקרב רובוטים, הם לא יהיו אוטונומיים אלא מופעלים מרחוק, בידי מפעיל שיתרכז בהפעלת הרובוט שלו, ואשר בגלל מגבלות התקשורת ברמה הטקטית הנמוכה לא יוכל להיות מרוחק ממנו. בכל מקרה, קשה, אם לא בלתי אפשרי, לממש רובוט לוחם רב־תכליתי. את רוב הלוחמים הנדרשים לניהול קרב המגע, לא יוכלו הרובוטים שיוכנסו לשימוש להחליף, ובמיוחד לא את הלוחמים בסמטאות של השטח הבנוי. רובוטים שיופעלו בקרב, בעתיד הנראה לעין, יופעלו רק בתפקידים מוגדרים ומצומצמים, ויותאמו בדיוק למשימתם.

רשימת מקורות

- כהן-אינגר, נורית וגל א' קמינקא. "והרי התחזית: צה"ל בדרך לצבא תבוגי - מפת דרכים לאימוץ טכנולוגיות בינה מלאכותית בצה"ל", **בין הקטבים** 18, דצמבר 2018. עמ' 91-118.
- כהן-אינגר, נורית, וליאור זכרון. "לראות את הנולד" - שילוב יכולות בינה מלאכותית לשיפור העוצמה האווירית של צה"ל", **בין הקטבים** 11-12, 2016. עמ' 229-246.
- עקביה, גדעון. "שו"ב ממוכן בכוחות השדה - בעיות יסוד". **מערכות** 407, יוני 2006. עמ' 17-27.
- שורש, דוד. "אסטרטגיה בעידן הבינה המלאכותית: תהליך תכנון אסטרטגי למגמות טכנולוגיות משבשות", **בין הקטבים** 18, דצמבר 2018. עמ' 119-148.
- Barrett, Brian. "An Artist Used 99 Phones to Fake a Google Maps Traffic Jam". *Wired*.
- Becker, Jeff. "How Lethal, Mobile, Protected and Aware? Exploring the Art of the Possible in Future Infantry Combat". *Small War Journal*, (8 Nov. 2019).
- Beede, Emma. "Healthcare AI systems that put people at the center". *Blog.google*, (25 Apr. 2020).
- Beede, Emma, Baylor, Elizabeth Elliott, Hersch, Fred, Iurchenko, Anna, Wilcox, Lauren, Ruamviboonsuk, Paisan and Vardoulakis, Laura M. "A Human-Centered Evaluation of a Deep Learning System Deployed in Clinics for the Detection of Diabetic Retinopathy". Paper presented at the CHI: Conference on Human Factors in Computing Systems.
- Benjamin, Wilson, Judy, Hoffman, and Jamie, Morgenstern. "Predictive Inequity" in *Object Detection*. *arXiv*, (21 Feb. 2019). :1902.11097v1 [cs.CV].
- Bloch, Sam. "Strawberry farmers: Come with Harvest CROO if you want to live". *The Counter*, (22 Jan. 2019).
- Brin, Sergey and Lawrence, Page. "The anatomy of a large-scale hypertextual Web search engine". *Computer Networks and ISDN Systems* 30, (1998), pp. 107-117.
- Broussard, Meredith. *Artificial unintelligence: how computers misunderstand the world*. MIT Press, (2018).
- Brown, Peter F., Cocke, John, Della Pietra, Stephen A., Della Pietra, Vincent J., Jelinek, Fredrick, Lafferty, John D., Mercer, Robert L. and Roossin Paul S. (June 1990). "A Statistical Approach to Machine Translation". *Computational Linguistics* 16, no. 2, (June 1990), pp. 79-85.
- Buolamwini, Joy and Gebru, Timnit. "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification". *Proceedings of Machine Learning Research* 81.
- Economist. "Technology in China". *The Economist Technology Quarterly* (4 Jan. 2020).
- Evtimov, Ivan, Eykholt, Fernandes, Earlenice, Kevin, Tadayoshi, Kohno, Atul, Prakash, Bo, Li, Rahmati, Amir and Song, Dawn. "Robust Physical-World Attacks on Deep Learning Models". *arXiv:1707.08945v4 [cs.CR]*, (13 Sep. 2017).
- Eykholt, Kevin, Evtimov, Ivan, Fernandes, Earlenice, Li, Bo, Rahmati, Amir, Xiao, Chaowei, Prakash, Atul, Kohno, Tadayoshi and Song, Dawn. "Robust Physical-World Attacks on Deep Learning Visual Classification". *Conference on Computer Vision and Pattern Recognition* (10 Apr. 2018).
- Goodfellow, Ian, Bengio, Yoshua and Courville, Aaron. *Deep Learning*. MIT Press (2016), p.19.

- Google. "Search Quality Evaluator" (5 Dec. 2019).
- Halevy, Alon, Norvig, Peter and Pereira, Fernando. "The Unreasonable Effectiveness of Data". *IEEE Intelligent Systems* 24 (Mar-Apr 2009), no. 2, pp. 8-12.
- Hao, Karen. "Doctors are using AI to triage covid-19 patients. The tools may be here to stay". *Technology Review*. (23 Apr. 2020)
- Hofstadter, Douglas R. *Le ton beau de Marot: in praise of the music of language*. Basic, (1997).
- Hofstadter, Douglas. "The Shallowness of Google Translate". *The Atlantic*. (30 Jan. 2018).
- Hutson, Matthew. "Hackers easily fool artificial intelligences - Adversarial attacks highlight lack of security in machine learning algorithms". *Science* 361 (20 July 2018), no. 6399.
- Knight, Will. "Facebook's Head of AI Says the Field Will Soon 'Hit the Wall'". *Wired* (12 Apr. 2019).
- Lawrence, Page, Brin, Sergey, Motwani, Rajeev and Winograd, Terry. "The PageRank Citation Ranking: Bringing Order to the Web". (29 Jan. 1998).
- Lee, Timothy B. "New Data Backs Up Details in a Fatal 2018 Tesla Model X Crash," *Wired* (18 Feb. 2020).
- LeCun, Yann, Bengio, Yoshua. and Hinton, Geoffrey. "Deep learning: Review". *Nature* 521 (28 May 2015), pp. 436-444.
- Lewis-Kraus, Gideon. "The Great A.I. Awakening: How Google used artificial intelligence to transform Google Translate". *New York Times* (14 Dec. 2016).
- Liu, Evie. "Quant Funds Are Trailing the Traditionalists in Coronavirus Crisis". *Barrons.com* (5 Apr. 2020).
- Lohr, Steve. "At Tech's Leading Edge, Worry About a Concentration of Power". *New York Times* (26 Sept. 2019).
- McCallum, Andrew. "Energy and Policy Considerations for Deep Learning in NLP". *arXiv:1906.02243v1 [cs.CL]* (5 June 2019).
- McDermott, Drew. "Artificial Intelligence Meets Natural Stupidity" *SIGART* (Apr. 1976).
- Metz, Cade. "A.I. Is Learning from Humans. Many Humans: Artificial intelligence is being taught by thousands of office workers around the world. It is not exactly futuristic work". *New York Times* (16 Aug. 2019).
- National Transportation Safety Board. "Tesla Crash Investigation Yields 9 NTSB Safety Recommendations", *News Release* (25 Feb. 2020).
- Obermeyer, Ziad, Powers, Brian, Vogeli, Christine and Mullainathan, Sendhil. "Dissecting racial bias in an algorithm used to manage the health of populations". *Science* 366 (25 Oct. 2019).
- O'Neil, C. *Weapons of math destruction: how big data increases inequality and threatens democracy*. Broadway (2016), pp.137-138.
- Rosenblatt, Frank. "The perceptron: A probabilistic model for information storage and organization in the brain". *Psychological Review* 65 (1958), pp. 386-408.
- Rumelhart, David E., Hinton, Geoffrey E. and Williams, Ronald J. "Learning representations by back-propagating errors". *Nature* 323 (9 Oct. 1986). pp. 533-536.

- Schwartz, Roy, Dodge, Jesse, Smith, Noah A., Etzioni, Oren Strubell, Emma and Ananya, Ganesh. *Green AI*. (Allen Institute) (July 2019).
- Seabrook, John. "The Age of Robot Farmers". *The New Yorker* (15 Apr. 2015).
- Sheikh, Knvul. "A Growing Presence on the Farm: Robots". *New York Times* (13 Feb. 2020).
- Smith, Brad. "Facial recognition technology: The need for public regulation and corporate responsibility". *Microsoft Blog* (6 Dec. 2018).
- Szabo, Liz. "Artificial Intelligence Is Rushing into Patient Care - And Could Raise Risks". *Scientific American* (24 Dec. 2019).
- Thornhill, John. "Trusting AI too much can be fatal". *Financial Times* (3 March 2020).
- Thys, S., Van Ranst, W. and Goedemé, T. "Fooling automated surveillance cameras: adversarial patches to attack person detection". *arXiv:1904.08653* (2019).
- Tschandl, Philipp et al. "Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study". *The Lancet: Oncology* 20 (1 July 2019), no. 7, pp. 938-947.
- Vinyals, Oriol, Toshev, Alexander, Bengio, Samy and Erhan, Dumitru. "Show and Tell: A Neural Image Caption Generator" (20 Apr. 2015). *arXiv:1411.4555v2 [cs.CV]*.