

## זיהוי תנועה במוח האדם מתרחש מהר מאוד

Andrea Tacchetti<sup>1</sup>, Leyla Isik<sup>2,3\*</sup>

<sup>1</sup>DeepMind טכנולוגיות בעירבון מוגבל, לונדון, בריטניה

<sup>2</sup>המחלקה למדעי המוח והקוגניציה, קיימברידג', מסצ'וסטס, ארצות הברית

<sup>3</sup>המחלקה למדעי הקוגניציה, אוניברסיטת ג'ונס הופקינס, בלטימור, מרילנד, ארצות הברית

"מה האדם הזה עושה?" בלי בכלל לחשוב על כך, המוח שלכם מעבד את מה שאתם רואים, ומשתמש במידע כדי לזהות את הפעולות של אנשים אחרים, כמו ריצה והליכה. זיהוי הפעולות הוא פשוט עבור בני אדם, אולם הוא בעיה מאתגרת מאוד עבור מחשבים. בפרט, קשה למחשבים לזהות את אותן הפעולות מנקודות מבט שונות, כמו ממבט קדימה או מהצד, או כנגד רקעים שונים. כתוצאה מכך, מדענים ומהנדסים לא היו מסוגלים לפתח אלגוריתמים של מחשבים שיכולים לזהות פעולות טוב כמו בני אדם. מטרת המחקר שלנו היא להבין טוב יותר כיצד מוחות של אנשים מזהים פעולות כך שנוכל לשכפל את התפקודים האלה במערכות מחשב.

באמצעות MEG, מכשיר שמאפשר לנו לקרוא מוחות ואותות, ולמידת מכונה, ענף במדעי המחשב, אנו מסוגלים לפענח את האותות שמתרחשים במוח האדם כשהוא מזהה פעולה. מאחר ש-MEG מאתר שינויים מהירים באותות מוחיים, הוא מאפשר לנו להבין כיצד מידע על פעולה מתגבש במהלך הזמן. מצאנו שהמוח מזהה פעולות מהר מאוד, בתוך 200 מילישניות (מילישניה אחת היא 1/1000 שנייה). יתרה מזו, המוח מתעלם מיד משינויים בנקודת המבט, כמו מבט מקדימה או מבט מהצד, ועדיין מזהה פעולה באותה המהירות. באמצעות הבנת

### סוקרים צעירים

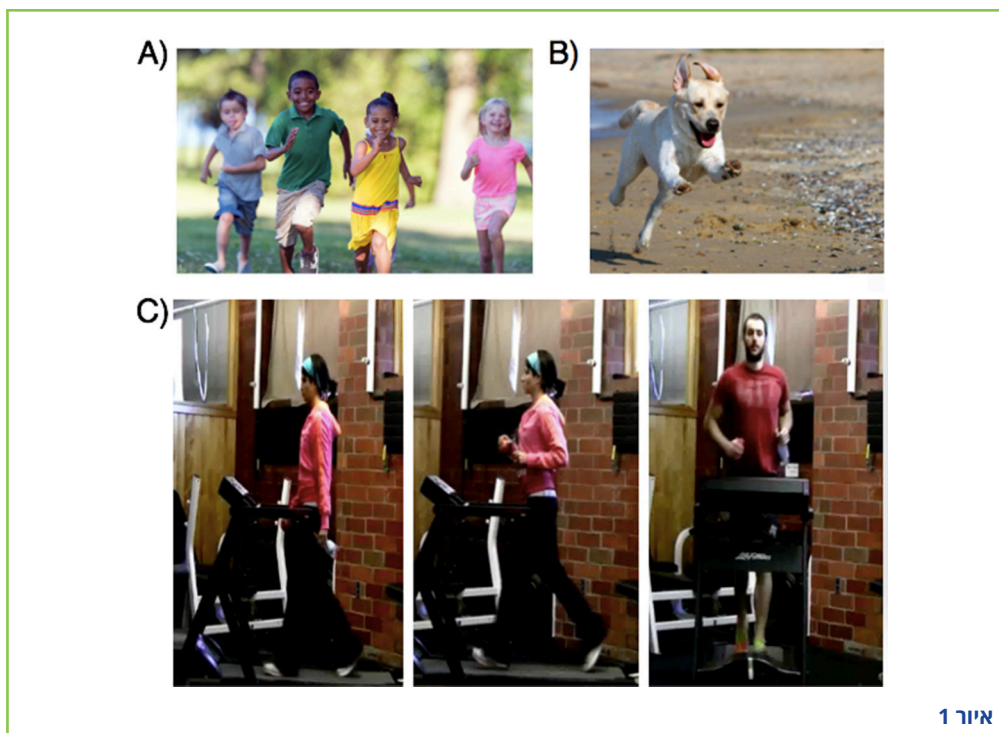
DUKE  
NEUROSCIENCE  
CAMP 2018



גיל: 9-11

## איור 1

(A,B) אנשים יכולים לזהות בקלות פעולות שונות כמו ריצה, ללא תלות במי שמבצע אותן; הרקע של הסצנה או נקודת המבט. (C) תמונות סטילס מווידיאו שהשתמשו בהן בניסוי שלנו שבהן אנשים מבצעים פעולות שונות מנקודות מבט שונות. האתגר למחשבים הוא לזהות פעולות דומות מאוד – כמו אותו אדם שהולך לעומת רץ, וגם לזהות את אותה הפעולה שמבוצעת על-ידי אנשים שונים מנקודות מבט שונות.



איור 1

הסדר שבו המוח מפרק את הבעיה המורכבת הזו, למדנו דברים רבים על חישובים מוחיים שמעורבים בזיהוי פעולות אנושיות.

## טרנספורמציות

### (Transformations)

שינויים בתמונה אשר משנים את האופן שבו הסצנה נראית, כמו שינויים בנקודת המבט, אולם אינם משנים באופן משמעותי את התכנים שלה.

## הבחנה

### (Discrimination)

היכולת להבחין בין דברים שונים (כמו זיהוי ריצה לעומת הליכה במקרה של פעולות).

## הכללה

### (Generalization)

היכולת לזהות משהו למרות שינויים או טרנספורמציות חזותיות.

## אלגוריתם

### (Algorithm)

אוסף של חוקים או חישובים שהמחשב מבצע כדי לפתור בעיה.

## הבעיה

זיהוי הפעולות של אנשים אחרים הכרחי בחיי היום-יום שלנו. אנו יכולים לומר בקלות כשמישהו מנופף לנו או רץ הרחק ממהשו. זה מסייע לנו להחליט מה לעשות הלאה. זיהוי הפעילות גם נהיה חשוב יותר ויותר עבור מערכות ממוחשבות כדי לזהות פעולות בזמן שמדענים מפתחים בינה מלאכותית (AI) שצריכה לתקשר עם בני אדם, כמו רובוטים מסייעים ומכוניות אוטונומיות.

באופן מרהיב, בני אדם מזהים פעולות אפילו כשישנן **טרנספורמציות** מורכבות: תמונה משתנה משנה את האופן שבו הסצנה נראית, אולם אינה משנה באופן משמעותי את התכנים שלה. באיור 1, לדוגמה, קל לזהות את הפעולות בכל תמונה, למרות השינויים הרבים: המשתתפים שונים (ילדים, כלב ומבוגרים); מיקומי מצלמה שונים (לפני המשתתפים או מצידם), ורקעים שונים (פארק, חוף הים, חדר עם קיר לבנים).

בפרט, קחו את שלוש התמונות בשורה למטה באיור 1. אם תשוו אותן, שתי התמונות הראשונות נראות דומות יותר זו לזו מאשר שתי התמונות השניות. אולם, למרות הדמיון החזותי שלהן, אתם יכולים בקלות לומר ששתי התמונות הראשונות מציגות פעולות שונות (ריצה לעומת הליכה). אנו קוראים לזה "**הבחנה**". אתם יכולים גם לומר ששתי התמונות השניות כוללות את אותה הפעולה (ריצה), למרות השוני בין התמונות. אנו קוראים לזה "**הכללה**". היכולת גם להבחין וגם ולהכליל קשה אפילו עבור האלגוריתמים המתקדמים ביותר כיום. **אלגוריתם** הוא אוסף של חוקים או חישובים שמחשב מבצע כדי לפתור בעיות. הבחנה והכללה מבוצעות על-ידי תהליכים שונים במוח. כיצד הצעדים השונים האלה מפורקים על-ידי

המוח כדי לזהות פעולות? השָׁבָה על השאלה הזו תסייע לנו לזהות את הדרך המיטבית להטמיע את הצעדים האלה באלגוריתמים ממוחשבים, ולפתח בינה מלאכותית טובה יותר.

במחקר הזה שאלנו אם המוח האנושי מזהה פעולות שונות, ומתי הוא עושה זאת באופן שמכליל טרנספורמציות שונות. האם המוח מזהה ראשית פעולות ואז מקשר בין אותה הפעולה בנקודות מבט שונות? או האם הבחנה והכללה נפתרות באותו הזמן? באמצעות ידיעת מה המָחֵשֵׁב מְחַשֵׁב ומתי, נוכל להבין טוב יותר את התהליכים הבסיסיים שמתרחשים.

## מה עשינו וכיצד עשינו זאת?

### אוסף נתוני וידאו עבור זיהוי פעולה

עבור המחקר שלנו צילמנו סרטוני וידאו של אנשים שונים שמבצעים חמש פעולות יומיומיות שכוחות: ריצה; הליכה; קפיצה; אכילה ושתייה. צילמנו את הווידאו משתי נקודות מבט שונות: מבט מקדימה ומבט מהצד (ראו איור 1, שורה תחתונה). זה אפשר לנו לשאול אם האותות המוחיים מסוגלים להבחין בין פעולות שונות ולהכליל בין נקודות מבט שונות. שמרנו על שאר האלמנטים בווידאו כמו התאורה והרקע, קבועים. לאחר מכן היינו צריכים שיטה למדידת התגובה המוחית לסרטוני הווידאו השונים האלה.

### "קריאה מוחית" באמצעות טכניקה שנקראת מגנטואנצפלוגרפיה (MEG)

המוח האנושי מעבד המון מידע כל הזמן. אתגר עיקרי בחקר המוח האנושי הוא קריאת אותות מוחיים באופן שאינו פולשני: בלי לפתוח את גולגולתו של האדם ולרשום מתוך המוח שלו.

במחקר שלנו השתמשנו בכלי שנקרא מגנטואנצפלוגרף או MEG (איור 2). MEG מאתר שינויים בשדות המגנטיים, וזו הסיבה לכך שהוא נקרא "מגנטו"-אנצפלוגרפיה. כאשר מיליוני תאי עָצֵב (תאים במוח) מופעלים, או "יורים" יחד, הם מייצרים שינויים בשדה המגנטי שמסביב לראש. השינויים החלשים מאוד האלה הרבה יותר חלשים מהשדה המגנטי של כדור הארץ, אולם אנו עדיין יכולים לאתר אותם באמצעות MEG מאחר שהוא רגיש מאוד.

ל-MEG יש קסדה עם 306 חיישנים שמודדים את השדה המגנטי החלש של המוח. שלא כמו שיטות אחרות של מדידת פעילות מוחית, MEG מאפשר לנו לאסוף הרבה מידע במהירות רבה: אנו יכולים לקבל מדידת MEG חדשה 1,000 פעמים בשנייה. המחיר הוא שאיננו יכולים לומר בדיוק מאיזה אזור במוח האותות מגיעים. כתוצאה מכך, אנו משתמשים באותות מכל המוח. האותות האלה מורכבים מאוד, ואנו זקוקים לדרכים חכמות כדי לנתח אותם.

אנו משתמשים ב**למידת מכונה** כדי לאתר את נתוני ה-MEG שלנו (ראו תיבה 1). אנו עושים זאת באמצעות שימוש ב-MEG על אנשים רבים בעודם צופים בסרטוני וידאו של פעולות שציינו קודם לכן, כך שהאלגוריתם הממוחשב יכול ללמוד כיצד נראית התבנית של הפעילות המוחית בתגובה לכל פעולה. זה נקרא **אימון**. לאחר מכן, ברגע שהאלגוריתם למד, הצגנו תבניות חדשות שהאלגוריתם לא ראה לפני כן, מה שנקרא **בחינה**, ושאלנו את האלגוריתם "באיזה וידאו האדם הזה צפה?" במילים אחרות, איזו תבנית מנתוני האימון שלנו הכי דומה

#### למידת מכונה (Machine learning)

ענף של מדעי המחשב שמטרתו לגרום למחשבים ללמוד מטלות מורכבות שמבוססות על נתונים.

#### אימון (Training)

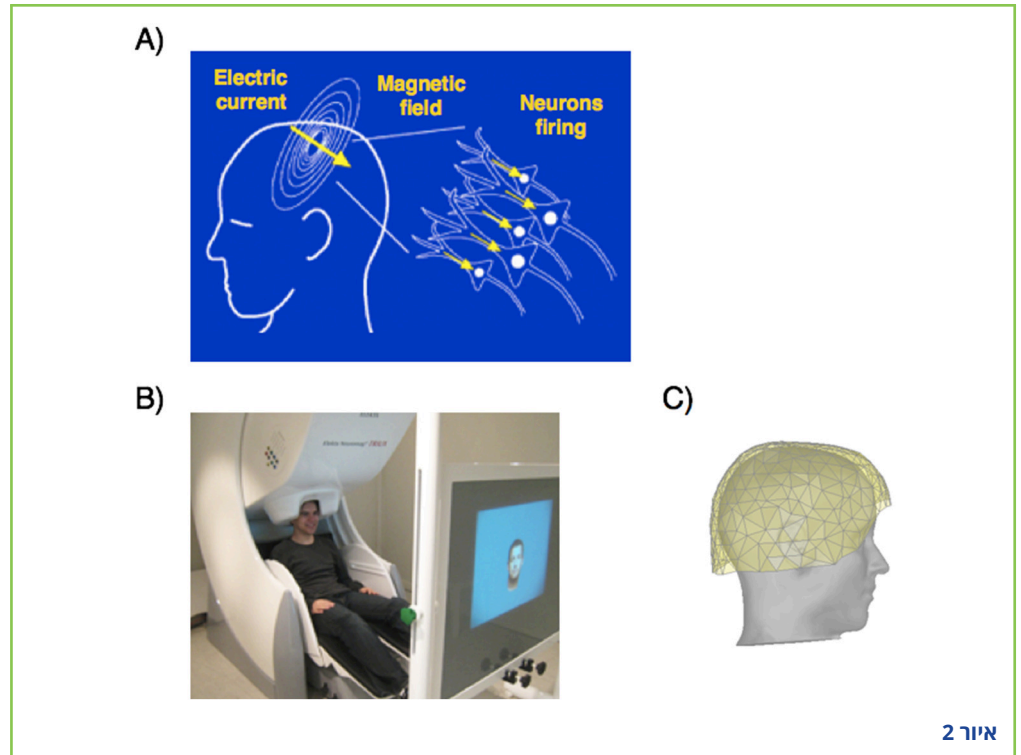
תהליך למידת המכונה שמשתמשים בו במטרה למצוא מסווג (ראו הגדרה בהמשך) עם קבוצה אחת של נתונים.

#### בחינה (Testing)

תהליך למידת המכונה שנועד להעריך את המסווג (ראו הגדרה בהמשך) עם קבוצה חדשה של נתונים (שלא השתמשו בה במהלך האימון).

**איור 2**

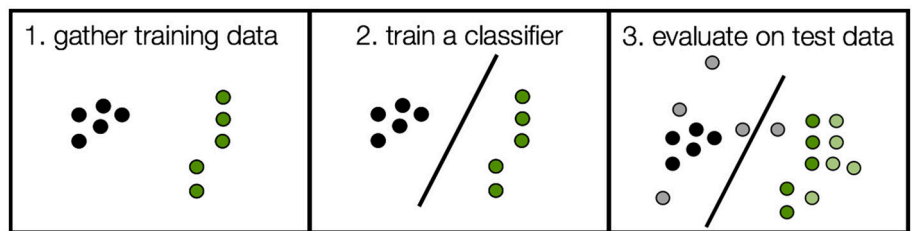
(A) מדידות MEG של פעילות מוחית באמצעות איתור השינויים בשדה המגנטי שמויצר על-ידי מיליוני תאי עצב (נוירונים) ש"יורים" באותו הזמן. תאי העצב האלה "יורים" ומייצרים זרם חשמלי, מה שמוביל לשינויים בשדה המגנטי. (B) אדם שיושב בתוך מכונת MEG אשר מכילה חיישנים בקסדה שסביב ראשו של האדם. (C) ציור של קסדת MEG אשר מכילה 306 חיישנים שמפוזרים באופן זהה ברחבי הראש. כל חיישן מודד את השינוי בשדה המגנטי שסביבו.



איור 2

למה שמופיע בנתוני הבחינה? לפעמים זה נקרא "פענוח עצבי". זה מאפשר לנו לקרוא באיזו פעולה האדם צפה, מתוך הפעילות המוחית שלו. אנו משתמשים בדיוק של הקריאה כמדד לכך שהאותות המוחיים מסוגלים להבחין בין פעולות שונות.

**תיבה 1 מהי למידת מכונה?**



**למידת מכונה** היא ענף במדעי המחשב שמטרתו ללמד מחשבים לפתור מטלות מורכבות שמבוססות על נתונים. לדוגמה, אנו יכולים לתת למחשב נקודות ירוקות ושחורות - הנתונים האלה נקראים **נתוני אימון**. אנו יכולים לבקש מהמחשב למצוא את הקו שמפריד את הנקודות בצורה הטובה ביותר. הקו הזה נקרא **מסווג**. לאחר מכן אנו יכולים להשתמש בו כדי "לסווג" את הנקודות האלה כירוקות או שחורות. נקודות הנתונים החדשות האלה נקראות **נתוני בחינה**; (הצבעים תואמים לצבעים של נתוני האימון, אולם הם מודגשים בגוון בהיר יותר), והמחשב מעולם לא ראה אותן קודם. בתמונה שלעיל הם מוצגים כנקודות הבהירות יותר בפאנל 3. אנו יכולים לראות כמה הקו המסווג מדויק באמצעות כמות נקודות הבחינה שהוא סיווג נכון. במקרה שלנו, זה יוצא 9 מתוך 10 או 90%. אם המסווג היה מבצע ניחושים אקראיים הסיווג היה מצליח באופן מקרי. במקרה כזה הסיכוי היה 50% (1 מתוך 2), מאחר שישנן שתי קטגוריות שונות.

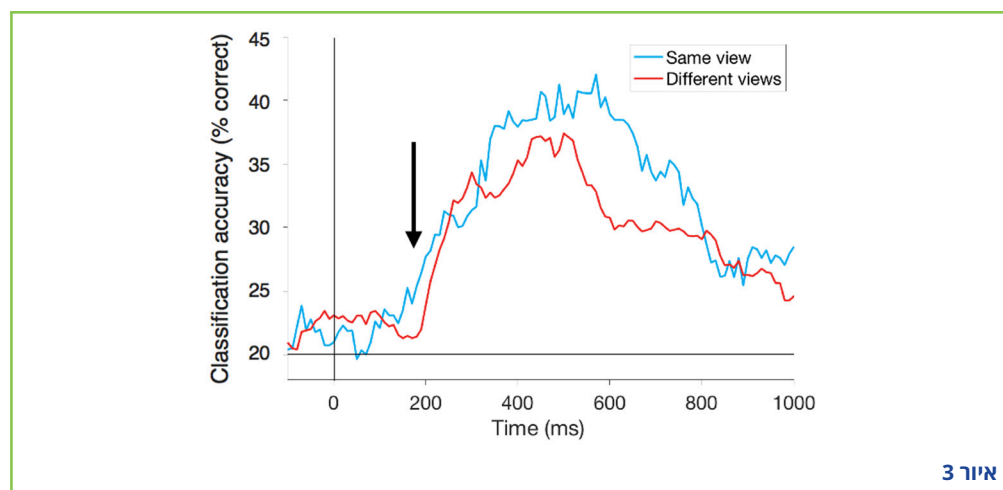
בניסוי שמתואר למטה, במקום ירוק לעומת שחור המסווג מאומן להבחין בין חמש פעולות שונות. הסיכוי האקראי לצדוק במקרה הזה הוא 20% (1 מתוך 5). אנו חוזרים על תהליך האימון והבחינה הזה בכל נקודה בזמן כדי לקבל מדד לדיוק הסיווג במשך הזמן. אנו משתמשים בדיוק הסיווג כמדד להאם אותות מוחיים מסוגלים להבחין בין קטגוריות שונות (ירוק לעומת שחור, או ריצה לעומת הליכה).

**מסווג (Classifier)**

אלגוריתם של למידת מכונה שמשתמשים בו כדי לפתור מטלות באמצעות חלוקת הנתונים לקטגוריות שונות.

### איור 3

קריאת מידע על פעולה מהמוח האנושי. העקומות מראות את הדיוק של המסווג (אחוז הניסיונות שסווגו נכון) בקריאת הפעולה שלנו מהפעילות המוחית של המשתתף (ציר y) במהלך הזמן (ציר x). זמן 0 מציין מתי הוודאו מתחיל. מ-200 מילישניות אחרי שהוודאו מתחיל (חץ שחור) אנו מסוגלים לומר באיזו פעולה האדם צופה ברמת דיוק שהיא באופן מובהק מעל לרמת המקרה. זה נכון אם אנו מסתכלים על פעולות שצולמו מאותה נקודת המבט, כמו במבט מקדימה (בכחול) ואפילו בין נקודות מבט שונות (באדום).



איור 3

אנו חוזרים על התהליך הזה בכל פעם שאנו מקבלים קריאת MEG חדשה (פעם אחת כל מילישנייה). זה מאפשר לנו לראות כיצד מידע במוח משתנה עם הזמן.

## מה מצאנו?

### המוח האנושי מזהה פעולות מהר

משתתפי הניסוי ישבו ב-MEG וצפו בסרטוני הוודאו. אנו השתמשנו במערכת למידת המכונה שלנו בשביל לצפות, בכל נקודה בזמן, איזו פעולה הצופים ראו – ריצה; הליכה; קפיצה; אכילה או שתיה. אם סיווג האלגוריתם היה מדויק מעל לאחוז מקרי משמעות הדבר שהאותות המוחיים שרשמנו כללו ידע שיכול להבחין בין פעולות.

כדי להתחיל הסתכלנו על סרטוני הוודאו שצולמו כולם מאותה נקודת המבט (לדוגמה במבט מקדימה). לפני שהוודאו מתחיל בזמן 0 הניבוי שלנו הוא בערך 20% (או 1 מתוך 5), כמו של סיכוי מקרי, מאחר שאין כלום על המסך (איור 3, עקומה כחולה) והמסווג מנחש באופן אקראי. ברגע שהוודאו מתחיל אנו רואים עלייה מובהקת בדיוק של המסווג, מה שמורה על כך שישנו מידע שקשור לפעולה אשר נוכח באותות המוחיים. פסגת הדיוק של המסווג היא בסביבות 40%. למטרות שלנו אנו מעוניינים לדעת מתי המידע על הפעולה מופיע לראשונה באותות ה-MEG, כך שנוכל להסתכל על הפעם הראשונה שהפענוח הוא מעל לרמת המקרה (בהתבסס על מבחנים סטטיסטיים). הדיוק של המסווג מגיע לרמה הזו בזמן של 200 מילישניות אחרי שהוודאו מתחיל. זה מהר מאוד, ואחרי שש תמונות בלבד מתוך הוודאו.

### המוח האנושי מזהה פעולות בין נקודות מבט שונות

בניסוי הקודם אלגוריתם למידת המכונה שלנו אָמַן ונבחן על נתונים ממשותתפים שצופים בפעולות בוודאו מנקודת מבט אחת, כמו למשל במבט מקדימה. בסופו של דבר אנו נרצה להבין מתי המוח יכול לבצע הכללה בין נקודות מבט שונות. כדי לעשות זאת, אימנו את האלגוריתם שלנו על נתוני אימון ממשותתפים שצפו בסרטוני הוודאו מנקודת מבט אחת, ובחנו את האלגוריתם על נקודת מבט אחרת.

ממש כמו בניסוי הקודם, לפני שהווידיאו התחיל בזמן 0, המסווג מנחש באופן אקראי (הדיוק במקרה כזה הוא 1 מתוך 5, או 20%). ברגע שהווידיאו מתחיל, אנו שוב מנבאים באיזו פעולה המשתתף צפה בזמן של 200 מילישניות אחרי שהווידיאו מתחיל (איור 3, עקומה אדומה). משמעות הדבר היא שאותות מוחיים יכולים לבצע הכללה בין נקודות מבט שונות. לא רק שהמוח מזהה במהירות פעולות שונות, אלא שהוא גם עושה זאת בדרך שמתעלמת משינויים בנקודת המבט. באופן מעניין, שתי המטרות ההפוכות האלה של הבחנה והכללה, מתרחשות באותו הזמן. אנו יודעים זאת מאחר שהעקומות הכחולה והאדומה נמצאות זו מעל זו, מה שמראה שברגע שהמוח מזהה את הפעולות הוא יכול להתעלם משינויים בנקודת המבט. הממצא הזה מציע שגם הבחנה וגם הכללה מתרחשות באותו השלב במוח, ולכן זה עשוי להיות הגיוני להטמיע

אותן באותו השלב באלגוריתם ממוחשב.

## למידת האופן שבו המוח האנושי מזהה פעולות יכולה לסייע לנו ליצור בינה מלאכותית יעילה יותר

זיהוי פעולה הוא היבט הכרחי בחיים היומיום של אנשים, וקישור הכרחי עבור מערכות של בינה מלאכותית כדי שיוכלו לתקשר עם בני אדם בזהירות ובאופן משמעותי. בעבודה שלנו זיהינו את האותות המוחיים שמבחינים בין שתי פעולות שונות, והראינו שהאותות האלה מתרחשים במהירות גדולה מאוד – אחרי שרואים 200 מילישניות של וידיאו בלבד. באופן חשוב, האותות המוחיים האלה חייבים להיות מספיק גמישים כדי לעבוד בתנאים שונים. הראינו שהאותות המוחיים האלה מיד מבצעים הכללה בין נקודות מבט שונות, גם כן לאחר 200 מילישניות בלבד.

המידע הזה על תזמון יכול לסייע לאלגוריתמים ממוחשבים לזהות פעולות בצורה טובה יותר, שדומה יותר לבני אדם. מהירות הזיהוי שראינו, לדוגמה, אומרת לנו שהאלגוריתם המוחי הוא יחסית פשוט, וככל הנראה מעבד את המידע החזותי פעם אחת בלבד (כלומר שלא צריך לזכור את הווידיאו כדי לזהות את הפעולה, או לעבד אותה כמה פעמים). העובדה שהמוח יכול לזהות באותה המהירות את אותן הפעולות מנקודות מבט שונות אומרת לנו שמערכות מחשב יכולות להתעלם מנתונים של נקודות מבט בשלבי העיבוד המוקדמים.

הבנת האופן שבו המוח מבצע זיהוי היא צעד ראשון בדרך להשלמת ההבנה של אינטליגנציה אנושית, והיא יכולה להדריך אותנו בפיתוח של מערכות טובות יותר של בינה מלאכותית.

## מאמר המקור

Isik, L., Tacchetti, A., and Poggio, T. 2018. A fast, invariant representation for human action in the visual system. *J. Neurophysiol.* 119:631–40. doi: 10.1152/jn.00642.2017

פורסם אונליין: 09 ביולי 2020

נערך על ידי: Michael A. Yassa, University of California, Irvine, United States



**ציטוט:** Tacchetti A and Isik L (2020) זיהוי תנועה במוח האדם מתרחש מהר מאוד. Front. Young Minds. doi: 10.3389/frym.2019.00009-he

### תורגם והותאם:

Tacchetti A and Isik L (2019) Fast action recognition in the human brain. Front. Young Minds 7:9. doi: 10.3389/frym.2019.00009

**הצהרת ניגוד אינטרסים:** המחברים מצהירים כי המחקר נערך בהעדר כל קשר מסחרי או פיננסי שיכול להתפרש כניגוד אינטרסים פוטנציאלי.

**COPYRIGHT** © 2019 © Tacchetti and Isik 2020. זהו מאמר בגישה פתוחה שמופץ תחת תנאי רישיון Creative Commons Attribution License (CC BY). השימוש, ההפצה או ההעתקה מותרים לשימוש בפורומים אחרים ובלבד שיינתן קרדיט למחבר(ים) המקוריים ולבעל זכויות היוצרים, ושהפרסום המקורי בעיתון זה מצוטט בהתאם למקובל באקדמיה. השימוש, ההפצה או ההעתקה אינם מותרים אם הם אינם עומדים בתנאים אלה.

## סוקרים צעירים

### DUKE NEUROSCIENCE CAMP 2018, גיל: 9-11

מחנה דייק למדעי המוח הוא חלק מתוכניות דייק לנוער. אנו מקבלים פניות מתלמידי בתי ספר בכל העולם. כל אחד יכול להצטרף, ולא נדרש נסיון קודם. הדבר היחיד שנדרש הוא עניין במדעי המוח! השנה דיברנו תכופות על מוחות ומכונות, ונהנינו כקבוצה מיישום מה שלמדנו על ממשקי מוח-מכונה ועל אינטראקציות אדם-רובוט כאשר סקרנו את המאמר של הדוקטורים Tacchetti ו-Islik.



## הכותבים

### ANDREA TACCHETTI

אני רוצה לבנות עולם שבו מחשבים ורובוטים יסייעו לנו לדאוג לעצמנו ואחד לשני. מחשבי עבר מסייעים לרופאים להבין מדוע אנו חולים; בקרוב הם ינהגו במכוניות שלנו ויסייעו לנו לטפל באנשים מבוגרים. אני עובד ב-DeepMind בלונדון היכן שאני בונה מערכות מחשב שמוסוגלות להבין מה אנשים ורובוטים אחרים עושים, ומה הם יעשו בהמשך. אם אצליח, מערכות אוטונומיות יפעלו באופן חלק בסביבה החברתית המורכבת שלנו.

### LEYLA ISIK

אני פרופסורית למדעי הקוגניציה באוניברסיטת ג'ונס הופקינס. אני מרותקת מהקישור בין אנשים לבין בינה מלאכותית (AI), ובפרט מהבנת הבעיות שקלות מאוד לאנשים (כמו למשל ראייה ותנועה בעולם שסביבנו), אולם הן הכי קשות עבור בינה מלאכותית. מטרת המחקר שלי היא להבין כיצד לזהות מידע על אנשים בהתבסס על מה שרואים. מחוץ למחקר אני אוהבת לרוץ, לבשל ולבלות זמן בחוץ. \*lisik@mit.edu



Hebrew version  
provided by

מחיאון המדע ע"ש בלומפילד ירושלים (ער.)  
متحف العلوم على اسم بلومفيلد القدس  
Bloomfield Science Museum Jerusalem

