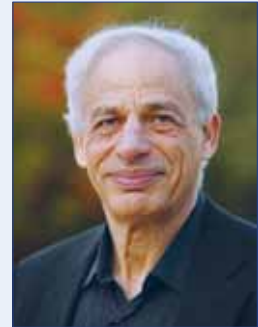


ראייה: מוח האדם ומכונות רואות



מאת פרופ' שמעון אולמן

פנים, הנמצאת בשימוש נרחב במצלמות. דוגמה אחרת היא מערכת ראייה ממוחשבת המשמשת עזר לנהג בהיותה מסוגלת לזהות מכוניות, הולכי רגל, נתיבי נסיעה ותמרורי תנועה ולהתריע על מצבים מסוכנים. מערכות מסוג זה יצאו לאחרונה מהמעבדה אל העולם המעשי וכבר מותקנות בדגמי רכב שונים. ראוי לציין כי מדינת ישראל היא מהמובילות בעולם ביישומי הראייה הממוחשבת, ותחום זה תורם לייצוא מערכות הנמכרות במאות מיליוני דולרים בשנה. בתור יצורים רואים אנחנו מודעים לתחושת הראייה, אבל מערכת הראייה במוח האנושי היא מערכת מורכבת שאנחנו מודעים לתוצאה הסופית של פעולתה אבל לא לדרכי פעולתה. דוגמה פשוטה לכך מוצגת באיור 1א. אנו רואים צורת משולש אף על



אנו מודעים לתוצאה הסופית של פעולת מערכת הראייה, אך לא לתהליכים המובילים לתוצאה זו.

1א. מערכת הראייה מסיקה קיומו של משולש בתמונה מעדויות חלקיות לקיומו.

1ב. אנו מקבלים מהתמונה מידע רב בסופם של תהליכים מורכבים במוח שטבעם ידוע חלקית בלבד.

ראייה באדם ובמכונה

תהליך הראייה הוא לכאורה טבעי וחסר מאמץ: כל אשר עלינו לעשות הוא לפקוח את עינינו ולהתבונן סביב. אולם למעשה זהו תהליך מורכב, שבחלקו הגדול עדיין אינו מובן לנו. מורכבותו של תהליך הראייה נסתר מאתנו מכיוון שרובו מתבצע ברמה בלתי מודעת של פעילות המוח. באופן מופלא המוח מסוגל להשתמש בקרינת אור המגיעה מהסביבה כדי ליצור הבנה מורכבת של העולם שמסביבנו, והבנת יכולת זו היא מטרתו של חקר הראייה.

בתחום זה של חקר הראייה בעיות המחקר שבהן אני עוסק נמצאות בשילוב של שני תחומים – מדעי המחשב ומדעי המוח. מבחינת מדעי המוח תהליך הראייה מתחיל בעין, אך עיקרו נעשה במוח. העין היא מעין מצלמה משוכללת הקולטת את התמונה ומעבירה אותה למוח, והמוח אחראי לעיבוד התמונה ולהבנתה. חלק ניכר מהמוח האנושי, בעיקר מהאזור העילי הנקרא קליפת המוח, מוקדש לתהליכי הראייה, ומסיבה זו בין היתר חקר הראייה הוא תחום מחקר מרכזי בחקר המוח. מבחינת מדעי המחשב קיים תחום מחקר פעיל הקרוי "ראייה ממוחשבת", אשר מטרתו ליצור מערכות שימושיות בעלות יכולת זיהוי עצמים והבנת תמונה, לפחות בתחומים מסוימים. מערכות כאלו כבר קיימות ונכנסו לשימוש. דוגמה מוכרת היא היכולת לאיתור



2א. תמונות של עצמים מאותה מחלקה, כגון "כלב", יכולות להיות שונות מאוד זו מזו, אך אנו מזהים אותן כשייכות לאותה מחלקה.

2ב. תמונה היא בבסיסה אוסף של נקודות שבכל אחת מהן נמדדת עצמת האור (בכמה אורכי גל). תהליך הכרת עצמים מתחיל בתמונה כזו בתור קלט ומזהה אוטומטית את העצם, או עצמים, המופיעים בה.

נמדדת עצמת האור (בכמה אורכי גל). הבעיה היא כעת לנתח אוסף זה של מדידות ולהבין מה יש בתמונה: חתול, כלב, מכונית או כל עצם אפשרי אחר. בנוסף, נרצה גם בתהליך הלומד מעצמו מאוסף של דוגמאות. כצופים אנושיים איננו מקבלים שיעור מפורש בהכרת עצמים אלא פשוט לומדים להכיר עצמים מדוגמאות הנקרות בדרכנו. אנו מעוניינים להבין את תהליך הלמידה ולשחזר תהליך דומה במערכות שימושיות. הבעיה המלאה העומדת בפנינו היא אפוא זו: ניתן לנו אוסף תמונות כגון אלו שבאיור 2 המכילות דוגמאות ממחלקה מסוימת, והמטרה היא ללמוד להכיר מחלקה זו באופן שיאפשר לזהות בעתיד כל עצם השייך לה.

הבעיה היא סבוכה, ולכן אגביל את הדיון שלהלן אך לכמה אלמנטים שעסקתי בהם ושהפכו לחלק מהגישה המקובלת במערכות ראייה ממוחשבת. הבסיס ללימוד הוא חיפוש אוטומטי באוסף הדוגמאות אחרי תבניות דומות של חלקים משותפים. לדוגמה, תמונות של פנים שונים בנויות

פי שהמשולש איננו מופיע במפורש בתמונה. מערכת הראייה מסיקה את קיומו מחלקי מידע המופיעים בתמונה, מבלי שיש לנו מגע ישיר ומודע עם תהליכי ההיסק שמפעילה המערכת. בתמונה מורכבת יותר, כגון זו שבאיור 1ב, אנחנו מקבלים בזמן קצר מידע עשיר על העולם שמסביב: אנחנו מבינים מה קרה, למשל היכן נמצאים האנשים בתמונה ומדוע. המחקר החישובי שבו אני עוסק מנסה להבין את תהליכי הראייה במוח המאפשרים הוצאת מידע מהתמונה וליישם במחשב.

זיהוי עצמים אוטומטי

ת וגמה לבעיה שבה עסקתי בשנים האחרונות היא הכרת עצמים. אתאר בקצרה להלן חלק קטן מעבודה זו. באופן טבעי אנחנו מחלקים את העולם שאנו רואים למחלקות של עצמים: אנו מתבוננים בעצם ומסווגים אותו כשייך למחלקת הכלבים, החתולים, המכוניות או לכל מחלקה אחרת. פעולת סיווג זו היא אתגר קשה בגלל השונות הגדולה שבין תמונות שונות של עצמים מאותה מחלקה. לדוגמה, תמונות של כלבים יכולות להיות שונות מאוד זו מזו ותלויות בסוג הכלב, בכיוון ההתבוננות, בתנאי התאורה ועוד (איור 2א). מערכת הראייה במוח מצטיינת במציאת המאפיינים שיאפשרו לנו לראות את המשותף לכל הכלבים למשל, ומה מפריד בינם לבין חתולים או לבין מחלקות עצמים אחרות. הבעיה של הכרת עצמים היא תחום מחקר עשיר בעל היסטוריה ארוכה; עשרות אלפי מאמרים הוקדשו לחקר בעיה זו במשך השנים.

כדי להבין את הקושי בהכרה אוטומטית של עצמים צריך לזכור שתהליך הראייה מתחיל בתמונה. תמונה המגיעה למחשב ממצלמה, או מגיעה למוח מהעין, אינה אלא אוסף מדידות של עצמות אור בשדה הראייה, כפי שמוצג באיור 2ב. כלומר, התמונה מיוצגת על ידי מספר רב של נקודות, ובכל נקודה

3

מציאת המרכיבים הבסיסיים של מחלקת עצמים

.א.



.ב.



	1	1	0	1	0	1	0	0
	1	0	0	1	1	1	0	0

.ג.



א3. אוסף תמונות לדוגמה מאותה מחלקה - סוסים במקרה זה. בתהליך לימוד אוטומטי נרכש אוסף המרכיבים הבסיסיים המאפיינים מחלקה זו.

ב3. תהליך זיהוי חלקים מאפיינים של מחלקת הסוסים. חלק מהתמונות שבשורה העליונה (מסומנות במסגרת ירוקה) מכילות סוסים. תמונות אלו מסומנות בספרה 1 בשורה שמתחת, ושאר התמונות מסומנות בספרה 0. בשורה התחתונה מופיע חלק מתמונת סוס המופיעה בחלק מהתמונות אך לא בכלן (המסומנות בספרה 1 בשורה התחתונה). החלק תורם לזיהוי ככל שהדמיון בין השורות גדול יותר. הדמיון נמדד באמצעות מדד מתמטי של "אינפורמציה הדדית".

ג3. אוסף מרכיבים מאפיינים של מחלקת הסוסים.

◀ כולן מאוסף מרכיבים משותפים, כגון עיניים, פה, גבות, קו שער וכדומה, הנמצאים ביחסים מרחביים דומים בתמונות השונות. גם כאשר נפגוש פנים חדשים, הם יהיו בדרך כלל שילוב אפשרי חדש של מרכיבים מוכרים מהעבר. מערכת המכירה את המרכיבים הבסיסיים ואת היחסים המרחביים ביניהם תוכל להכיר בהצלחה מגוון רחב של עצמים חדשים השייכים לאותה מחלקת עצמים. הקושי העיקרי הוא לזהות את אבני הבניין המשותפות של מחלקת עצמים, למשל של מחלקת כל הסוסים, שיפרידו בינה לבין כל מחלקה אחרת. הכיוון הבסיסי הוא אפוא לגלות באופן אוטומטי מאוסף דוגמאות את המרכיבים אשר שילובם מרכיב את מחלקת כל הסוסים, אך לא עצמים אחרים. כיצד אפשר אפוא לגלות באופן אוטומטי את ה"אל"ף-ב"ת" של תמונות כל הסוסים, כגון אלו המופיעים באיור 3א? כדי לענות על כך נעניין בעזרת איור 3ב בתמונות הנדרשות ממרכיב חלקי של מחלקת הסוסים. באיור 3ב התמונות שמכילות סוסים מסומנות בספרה 1, ואילו השאר מסומנות בספרה 0. מתחת לתמונות הסוסים מופיעה תמונה של מרכיב חלקי של סוס ואוסף כל התמונות שהוא מופיע בהן. התמונות שמכילות ביתר קירוב את המרכיב שבדוגמה מסומנות בספרה 1, ואילו השאר מסומנות בספרה 0. למרכיב תרומה נאה לזיהוי קיומה של התאמה בין השורות, כיוון שבמקרה כזה הימצאות החלק בתמונה היא עדות נאמנה לקיומו של סוס בתמונה. בפועל ההתאמה לעולם לא תהיה מלאה, ומשום כך יש צורך למדוד בשביל התאמה חלקית כלשהי, עד כמה החלק המוצא אכן תורם לזיהוי נכון. מתברר כי בחירת החלקים המתאימים לייצג מחלקה צריכה להיעשות באמצעות מדד מתמטי הקרוי מדד של "אינפורמציה הדדית". לא אכנס כאן לפרטים הטכניים של מדד זה, אך הרעיון הבסיסי הוא פשוט וטבעי: מחפשים חלקים המופיעים במספר גדול ככל האפשר של תמונות סוסים ובמספר קטן ככל האפשר בתמונות אחרות שאינן מכילות סוסים. השימוש

ייצוג עצמים במוח האנושי

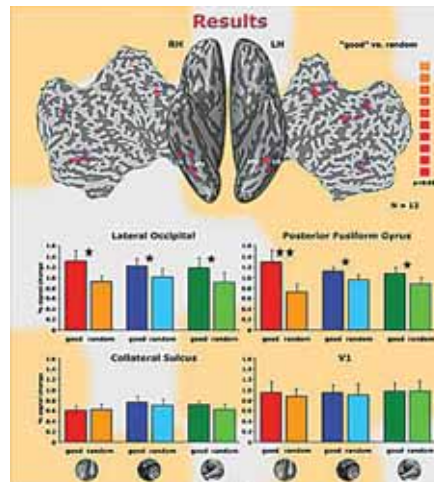
אינו כי התהליך המתמטי מביא לידי זיהוי עצמים אוטומטי במחשב. הבעיה העומדת בפני מערכת הראייה האנושית דומה בביססה לבעיה העומדת בפני מערכת ראייה ממוחשבת, ותהליך דומה למתואר לעיל ניתן ליישום באופן טבעי גם במנגנונים עצביים. אפשר בהחלט לשער כי גם המוח האנושי עשוי לייצג עצמים על ידי שמירה בזיכרון של תמונות של חלקים בסיסיים שנבחרו על סמך תרומתם האפשרית לזיהוי עצמים ממחלקות שונות. מחקרים שונים שנעשו בקבוצות שונות בעולם אכן תומכים בהשערה זו.

לדוגמה, מחקר שנעשה במעבדתו של פרופ' רפאל מלאך במכון ויצמן למדע בחן שאלה זו על ידי שימוש בהדמיה מוחית באמצעות מערכת הדמיה תפקודית של תהודה מגנטית [5]. בשיטה הזו סורקים מוח של נבדק בתוך מגנט רב עצמה

במדד של אינפורמציה הדדית מאומץ מכיוון שהוא מביא לידי צמצום האפשרות של טעות הזיהוי בתמונות חדשות. על יסוד זה אפשר לבחור בחירה אוטומטית של המרכיבים, שהם מעין "אל"ף-בי"ת" של מחלקת עצמים כלשהי. דוגמאות ממחלקת הסוסים מופיעות באיור 3.3. אוסף המרכיבים ומיקומם היחסי מהושך בהמשך בסיס לזיהוי סוסים כלשהם בתמונות חדשות. לדוגמה, כל הסוסים באיור 3.3 א זוהו בהצלחה לאחר למידה אוטומטית של מחלקת הסוסים למרות הגיוון הרב בין התמונות השונות. הדיון לעיל הוא כמובן חלקי ביותר. בפועל תהליכי הלימוד והזיהוי מורכבים הרבה יותר, אך תהליך מהסוג שתיארתי לעיל, בהרחבות שונות, הוא בסיס למערכות עכשוויות עם ביצועים טובים בזיהוי מגוון רחב של עצמים [1-4]. אם עד לפני שנים לא רבות זיהוי עצמים אוטומטי נחשב למטרה שאיננה ניתנת להשגה, כיום מערכות מסוג זה מתפתחות בקצב מהיר ומוכנסות לשימושים מעשיים בתחומים שונים.

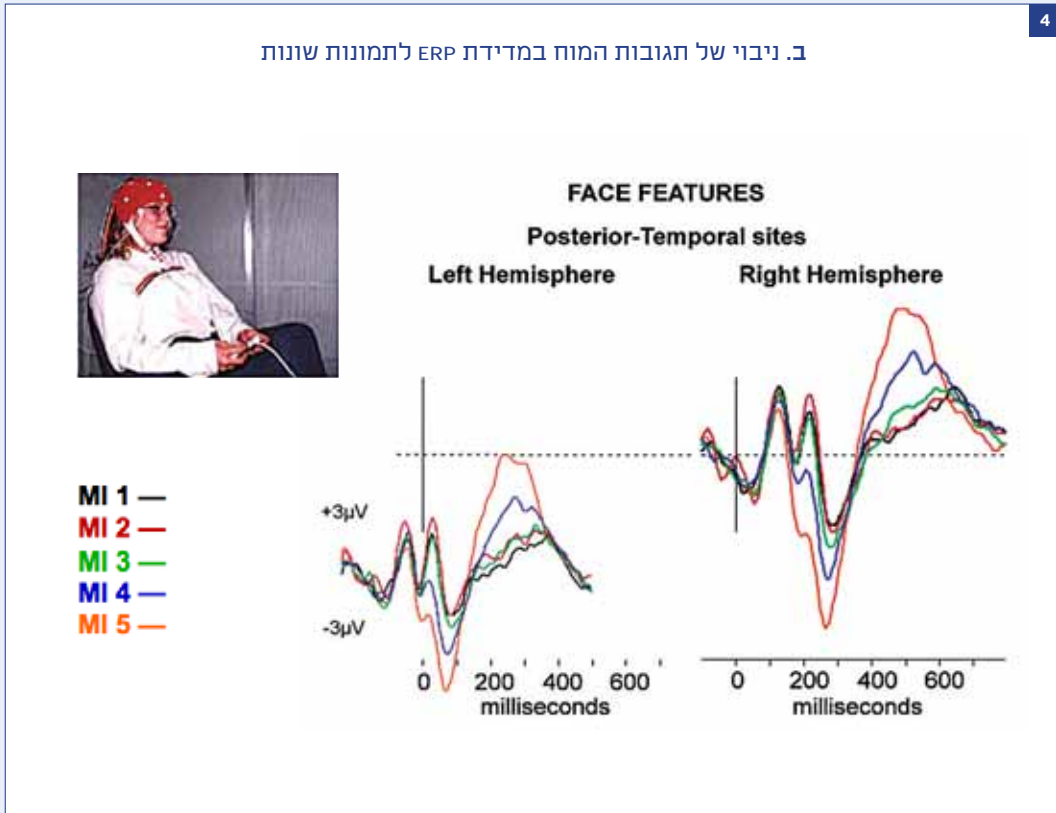
ייצוג במוח של מחלקות עצמים על ידי אוסף תמונות של מרכיבים בסיסיים

א. ניבוי תגובת אזורים במערכת הראייה וניתוחה



44. במהלך הניסוי נבדקים צופים במקטעים חלקיים של עצמים, כגון מכוניות או סוסים. פעילות המוח נמדדת באמצעות הדמיה תפקודית של תהודה מגנטית. פרטים נוספים בגוף המאמר. מתוך מחקר במעבדתו של ר' מלאך, מכון ויצמן למדע.

ב. ניבוי של תגובות המוח במדידת ERP לתמונות שונות



4. ניסוי המודד את פעילות המוח בזמן צפייה במקטעים חלקיים של תמונות עצמים שונים, שנמדדה באמצעות מדידת פעילות חשמלית במכשיר אלקטרו־אנצפלוגרפיה. הגרף השמאלי מציג מדידות מחציו השמאלי של המוח, והימני מחציו הימני. הגרפים מציגים את ההתאמה בין רמות פעילות יחסית הצפויות ממודל חישובי לבין מדידות שהתקבלו בפועל. פרטים נוספים בגוף המאמר. מתוך מחקר במעבדתו של ש' בנטין, האוניברסיטה העברית בירושלים.

תמונת סוס, דרך מדד האינפורמציה ההדדית הניתן להערכה כמותית. המחקר העלה כי ככל שהמדד האובייקטיבי גבוה יותר, כך פעילות המוח כתגובה לחלק זה בתמונה, באזורי מוח המשתתפים בזיהוי עצמים, חזקה יותר בממוצע. דוגמה לכך אפשר לראות באיור 4. בחלקו העליון מופיעה מפה של אזורי המוח, ועליה מסומנים (בצבע אדום) אזורי ממערכת הראייה שנבדקו בניסוי. בחלק התחתון מופיעות מדידות של עצמת הפעילות שנמדדה מהמוח באזורים שונים ובתגובה לגירויים שונים. התוצאה העיקרית היא כי בשני האזורים הקשורים לזיהוי עצמים (בשורה העליונה של קבוצת הגרפים)

ומקבלים תמונה של פעילות המוח. מחקרים רבים בתחום הראו כי מערכת הראייה מורכבת מקבוצה של אזורים הקשורים זה לזה, ובהם אזורים הקשורים לתהליכים של זיהוי עצמים. במהלך המחקר הוצגו לנבדקים תמונות של חלקי עצמים שונים, כגון פנים שונים, סוסים ומכוניות. לכל אחד מהחלקים אפשר לקבל מהמודל החישובי ניבוי של פעילות המוח באזורים שונים של מערכת הראייה. בעיקר אפשר לקבל על כל חלק מדד משוער בדבר תרומתו האפשרית לזיהוי עצמים. לדוגמה, אפשר להשוות זה לזה שני חלקים שונים מתמונת סוס ולמדוד באופן חישובי מי מהם תורם יותר לתהליך הזיהוי של

ידע על העולם באמצעות מערכת הראייה במהלך ההתפתחות. נושא זה הוא בראשית דרכו ולא ייסקר כאן בפירוט [7], אך צפוי לו להערכתי מקום מרכזי הן בחקר המחשבה האנושית והן בפיתוחן של מכונות רואות. ■

ביבליוגרפיה

1. S. Ullman, *High-level Vision: Object Recognition and Visual Cognition*. Cambridge, MA, 1996.
2. S. Ullman, M. Vidal-Naquet, and E. Sali, "Visual features of intermediate complexity and their use in classification", *Nature Neuroscience*, 5(7) 1–6, 2002.
3. S. Ullman, "Object recognition and segmentation by a fragment-based hierarchy", *Trends in Cognitive Sciences* (11)2, 58–64, 2006
4. B. Epshtein, I. Lifshitz, S. Ullman, "Image interpretation by a single bottom-up top-down cycle", *Proc. Nat. Acad. Sci.*, 105(38) 14298–14303, 2008.
5. Y. Lerner, S. Ullman, B. Epshtein, and R. Malach, "Class information predicts activation by object fragments in human object areas", *Journal of Cognitive Neuroscience*, 20(17) 1–18, 2008
6. A. Harel, S. Ullman, B. Epshtein, and S. Bentin, "Mutual information of image fragments predicts categorization in humans: Electrophysiological and behavioral evidence", *Vision Research*, 47, 2010–2020, 2007.
7. S. Ullman, D. Harari, N. Dorfman, "From simple innate biases to complex visual concepts", *Proc. Nat. Acad. Sci.*, 2012.

ולשלוש מחלקות עצמים שונות, ההבדלים בפעילות המוח התאימו במובהק לצפי המבוסס על מדד אובייקטיבי העולה מהמודל החישובי. ניסוי המבוסס על שיטת מדידה אחרת של פעילות המוח, שנעשה במעבדתו של פרופ' בנטיין באוניברסיטה העברית בירושלים [6], העלה מסקנות דומות (איור 4ב). כמו בניסוי הקודם, גם בניסוי זה הוצגו לנבדקים תמונות חלקיות של עצמים שונים. החלקים סווגו מראש, לפני תחילת הניסוי, לחמש רמות שונות: ברמה 5, הגבוהה ביותר, קובצו כל החלקים שתרומתם לתהליך הזיהוי צפויה, על פי המודל, להיות הגבוהה ביותר, וברמה 1 קובצו החלקים שתרומתם הצפויה היא הנמוכה מכל החלקים שנמדדו. כמו בניסוי הקודם, גם כאן התקבלה התאמה גבוהה בין פעילות המוח כפי שנמדדה בניסוי לבין הצפוי מהמודל החישובי.

מלבד התמיכה בהשערה מסוימת הנוגעת לייצוגם של עצמים במערכת הראייה, מחקרים אלו הם גם דוגמה לשימוש במודלים חישוביים ככלי מחקרי בהבנת המוח. הסיבוך העצום של המוח מציב אתגר קשה בדרך להבנת פעילותו. כיום אפשר – באמצעות שיטות מדידה חדשות – לאסוף מגוון של נתונים על פעילות המוח, אך מכאן ועד להבנת תהליכי החישוב המתחוללים בו וכיצד הם תורמים למחשבה ולהבנת העולם שסביבנו, הדרך ארוכה וקשה. נראה כי לשילוב רב-תחומי של מדידות ניסיוניות ומודלים חישוביים יכול להיות ערך רב בהתמודדות עם אתגר מרתק זה.

היכולת לזהות עצמים ולהבין תוכנן של תמונות היא תהליך נרכש. לתינוקות עדיין אין יכולת לזהות עצמים. העולם הראייתי סביבם הוא חסר ארגון ובלתי מובן, אולם מוטבעות במוחם יכולות למידה משוכללות, ועם הזמן הם מתחילים להבין את מה שהם רואים ולרכוש ידע על העולם. בתהליך של רכישת ידע על העולם יש לחוש הראייה תפקיד מרכזי, ולכן היבט מרכזי של הבעיה שבה אני עוסק – מלבד המערכת הבוגרת – הוא רכישת