



האח הגדול מזהה אותך



מי טוב יותר בזיהוי פנים, אדם או מחשב?

במרץ 2007 פרסם המכון הלאומי לסטנדרטים וטכנולוגיה (NIST — National Institute of Standards) של ארצות-הברית את תוצאות מבחן יצרני תוכנות זיהוי הפנים של שנת 2006 (FVRT — Face Recognition Vendor Test – קישור בסוף המדור). המבחן כלל, בין השאר, השוואה של יכולת הזיהוי של כמה אלגוריתמים (שיטות הניתנות למימוש כתוכנות מחשב) ליכולת אנושית. התוצאות: שלושה אלגוריתמים הצליחו לזהות פנים טוב יותר מאשר בני-האדם שהשתתפו בניסוי, בכל התנאים שנכללו בניסוי עבור רמות התאורה ועבור רמת ההפרדה (הרזולוציה) של תמונות הפנים.

זוהי, כנראה, הפעם הראשונה שבה בוצע מבחן השוואתי בין בני-אדם לבין מחשבים בתחום זה. עם זאת, ידועות התוצאות הטובות ביותר שהשיגו מחשבים בשנים קודמות, כך שאפשר גם להסיק כי 2006 היתה השנה שבה באמת הצליחו מחשבים להדביק ולעבור את הביצועים האנושיים. כדי להבהיר זאת, יש להסביר את השיטה שבה נבדקים ההישגים בזיהוי: NIST וגופים אחרים המשתתפים בפיתוח ובבחינה של תוכנות זיהוי פנים אספו מאגרים גדולים של תמונות, כך שעבור כל אדם במאגר קיימות כמה תמונות. באחד המאגרים, לדוגמה, יש 108,000 תמונות של 36,000 אנשים. כדי לבחון את יכולת הזיהוי, מוצגות למחשב (או לאדם המשתתף בניסוי) שתי תמונות שנבחרו מתוך המאגר. על המחשב לדווח עד כמה הוא "מאמין" כי התמונות הן תצלומים של אותו אדם. ניסוי זה דומה לשימוש בפועל כאשר משווים תמונה של אדם בלתי ידוע למאגר תמונות של אנשים שזהותם ידועה, בתקווה למצוא התאמה של התמונה הראשונה לאחת מהויות הידועות.

במבחנים כאלה נהוג למדוד את הסיכוי של שתי צורות שגיאה: הסוג הראשון הוא "קבלה שגויה" (false accept), כלומר מצב שבו המחשב יבצע זיהוי שגוי, וידווח כי שתי תמונות הן של אותו אדם אף שאין הדבר כך. הסוג האחר הוא



"דחייה שגויה" (false reject), כלומר מצב שבו המחשב אינו מצליח לגלות כי שתי התמונות שהוצגו לו הן של אותו אדם. בדיווחים רבים נדרשות תוכנות המחשב להגיע לרמת קבלה שגויה של כ-0.001: רק באלפית מהמקרים שבהם ידווח המחשב על זיהוי יתברר בסופו של דבר כי היה זה זיהוי שווא. קל לשלוט בפרמטרים של התוכנה כך שהמחשב יגיע לרמת שגיאות נמוכה עד כדי כך, אך לשליטה זו יש מחיר: כאשר מורידים את הסיכוי של טעויות "קבלה שגויה", עולה הסיכוי לטעויות "דחייה שגויה". אילו היה המחשב אנושי, היינו מפרשים זאת כאילו המחשב הופך להיות זהיר והססן, ולכן חושש מטעות עד כדי כך שהוא נמנע במקרים רבים מלדווח על זיהוי אפשרי.

בעזרת מושגים אלה אפשר להשוות את התקדמות התוכנה בשני העשורים האחרונים, כאשר ההשוואה מתבצעת בכל מקרה עבור רמת קבלה שגויה של כ-0.001: ב-1993 הגיעו האלגוריתמים הטובים ביותר לתוצאה מאכזבת של 79% דחייה שגויה: המחשב לא זיהה ארבעה מכל חמישה "חשודים". לאלגוריתמים אלו היתה מגבלה נוספת: הם דרשו עזרה אנושית. עבור כל תמונה היה צורך באדם שיצביע בתוך התמונה על מיקומי שתי העיניים. ב-1997 כבר הופיעו אלגוריתמים אוטומטיים לחלוטין, שהגיעו לרמת דחייה שגויה של 54%. ב-2002 הושג שיפור נוסף, עד ל-20%. במאי 2004 הודיעה ממשלת ארצות-הברית על "האתגר הגדול של זיהוי הפנים" (FRGC — face Recognition Grand Challenge). תכנית זו, שהסתיימה במרץ 2006, הופעלה כדי לקדם את טכנולוגיות זיהוי הפנים, והציבה מטרה שאפתנית: שיפור הביצועים בסדר גודל אחד. מטרה זו הושגה. במבחן של שנת 2006 הושגה תוצאה של 1%: 99% מהזיהויים הנדרשים אכן בוצעו. תוצאה זו הושגה אמנם רק בתנאים הטובים ביותר, של הפרדה (רזולוציה) גבוהה ושל תנאי צילום מבוקרים, אך עדיין זוהי התקדמות מרשימה. לשם השוואה, עבור תנאי צילום לא-מבוקרים (כמו אלו שמתקבלים ממצלמות אבטחה), ההישגים הטובים ביותר שדווחו היו 11% דחייה שגויה עבור רזולוציה גבוהה ביותר (שישה מיליון פיקסלים בצילום, שמכיוון שבוצע בתנאים לא-מבוקרים תפסו הפנים רק חלק קטן מתוכו, והמרחק הממוצע בין מרכזי שתי העיניים היה כ-190 פיקסלים) ו-13% עבור רזולוציה גבוהה (ארבעה מיליון פיקסלים בתמונה, 110 פיקסלים בין העיניים). יש להניח כי

גם עבור תנאים אלו ישתפרו התוצאות בזמן הקרוב. כדאי לשים לב כי מה שנקרא בניסוי "רזולוציה גבוהה ביותר" הוא כבר עכשיו הרזולוציה המקובלת במצלמות דיגיטליות רבות.

מי צריך זיהוי פנים?

באותו דו"ח מבחן מובאות גם תוצאות של שתי טכנולוגיות זיהוי אחרות: זיהוי לפי צילומי הקשתית (החלק הצבעוני של העין, המקיף את האישון), ולפי סריקת פנים תלת-ממדית (כיום קיימים מכשירים מסחריים המאפשרים סריקה כזו). טכנולוגיות אלו הגיעו להישגים דומים לאלו של זיהוי פנים לפי צילומים דו-ממדיים.

מדוע משקיעה ממשלת ארצות-הברית מאמץ ומשאבים כה רבים בטכנולוגיות זיהוי? רמז לתשובה נמצא בדף הראשון של דו"ח המבחן, ברשימת המממנים, הכוללת את ה-FBI, את ה-DNI (Director of National Intelligence) – הממונה בממשלת ארצות-הברית על קהילת המודיעין ואת מחלקת "ביטחון המולדת" (Homeland Security). קל לדמיין את האפשרויות. כיום מוצבות מצלמות המכסות חלקים רבים מהמרחבים הציבוריים בעולם: מצלמות אבטחה ממשלתיות ומשטריות, מצלמות למניעת גנבות ומצלמות שמפעילים גופי תיירות ואנשים פרטיים. אם קהילת המודיעין האמריקנית תאגור תמונות של חשודים כטרוריסטים, היא תוכל, בעזרת טכנולוגיה זו, לסרוק תמונות ממקורות רבים אלו, ולקבל התרעות אמינות כאשר חשודים אלו יופיעו במקום כלשהו בעולם. כמה פיגועים אפשר למנוע כך? כמה קרבנות אפשר להציל?

קל לחשוב על שימושים נוספים: ראינו כבר עד כמה חרדה האוכלוסייה כאשר אסיר מסוכן ברח, בגלל החשש שהוא יבצע מעשי אלימות נוספים. כאשר כל מקום שבו יעבור האסיר יהיה מכוסה במצלמות, והתמונות המצולמות בהן ייבדקו, דרכו חזרה אל הכלא תהיה מהירה יותר. הורים שילדם נעלם חלילה יכולים גם הם להיעזר ביכולתם הבלתי נלאית של מחשבים לסרוק תמונה אחר תמונה, מצלמה אחר מצלמה, עד שהבן האובד יימצא.

אפשר לבדוק לא רק תמונות שצולמו ברגע זה, אלא גם תמונות שצולמו ברגע כלשהו בעבר: יותר ויותר מצלמות מחוברות לאינטרנט, כך שבעלי המצלמות יכולים להחליט אם לתת את הגישה לכל דורש או רק למורשים לכך. ברגע



בהפגנה שאינו מסכים עם מטרותיה? האם הוא יכול לבדוק איפה נמצא באמת עובד שדיווח כי הוא חולה? שרת החינוך יולי תמיר מספרת כי כאשר למדה בתיכון, התחמקה מבית-הספר כדי להשתתף בצעדה, אך לרוע מזלה זיהה אותה מנהל בית-הספר בצילום שפורסם בעיתון ליד כתבה על אותה צעדה. האם מנהלי בתי-הספר של מחר יקבלו לשולחנם בכל בוקר דיווח על אתרי הבילוי שבהם נצפו תלמידיהם? האם הורים צריכים להיות מעוניינים במעקב אחר מעשיהם של ילדיהם בני העשרה? (יש לזכור כי לא פחות סביר שאותם נערים ונערות, שהם הראשונים לאמץ חידושים טכנולוגיים, ימצאו את הדרך לעקוב אחר חבריהם, הוריהם ומוריהם).

האח הקטן צופה בך

לא רק ממשלות מתעניינות בזיהוי פנים. אחת החברות שהגיעה להישגים גבוהים במבחן FVRT נקראת נֶבֶן (Neven)

שיכולת הגישה מושגת (בדרכים חוקיות או אחרות), אפשר להקליט כל תמונה המתקבלת מכל מצלמות הווידאו הנגישות. גם אם הילד האובד אינו נראה כרגע בשום מצלמה, אפשר לחפש אותו בתמונות מהעבר, וכך לנסות לפענח מה קרה לו ולאן נעלם. דוגמה אחרת: אם המשטרה מנסה לקשר חשוד לפשע שבוצע במקום מסוים, היא יכולה לסרוק מצלמות באותו אזור סמוך לזמן שבו בוצע הפשע, בתקווה למצוא את החשוד. היא גם יכולה לנסות להתאים את ההקלטות למאגר תמונות העבריינים שברשותה.

חלק מהקוראים אולי חשים צמרמורת בנקודה זו: אותה טכנולוגיה עלולה לשמש כדי להגביל מאוד את חירויות הפרט. מה ימנע את המשטרה מלהשתמש במאגר התמונות מתעודות הזהות ורישיונות הנהיגה? האם מעסיק יכול לחפש את פני עובדיו בין המשתתפים



(Vision). באוגוסט 2006 רכשה את נבן חברת אינטרנט המוכרת לכולנו – גוגל. ההסבר של גוגל לרכישה זו מופיע בקישור שבסוף הכתבה: גוגל מתכננת לצייד את פיקאסה, תוכנת ניהול אלבום התמונות האישי שלה, ביכולת לסווג תמונות לפי נושא (אנשים או נוף, למשל), ובעתיד גם לזהות את האנשים והמקומות המופיעים בהן. מי לא היה רוצה לחפש בארכיון שלו את כל התמונות שצולמו במקום מסוים, או את כל התמונות שבהן מופיעה האחיינית החמודה? הרכישה של גוגל היא רק דוגמה אחת: חברות אחרות שהגיעו להישגים גבוהים הן סמסונג וטושיבה, שגם להן יש ודאי רעיונות איך לשפר (?) את חיינו בעזרת זיהוי פנים. אפשרות אחת, שכבר הוצגה בכמה מעבדות בעולם: כדי להימנע מהמצב המביך שבו אינני זוכר מיהו האדם שקרא הרגע בשמי, ארכיב על משקפיי מצלמה מיניאטורית. המצלמה תהיה מחוברת למחשב שאליו הזנתי תמונות ושמות, ואזונייה קטנה המוחבאת במוט המשקפיים תלחש לי את השם ששכחתי. אם תמונתו של אדם זה אינה במאגר שלי, יפנה המחשב לאינטרנט, דרך חיבור אלחוטי, ויבדוק אם הפנים המסתוריות מופיעות במקום כלשהו לצד שמו של האדם. גם זה לא הצליח? איאלץ לשאול מיהו אותו מכר ותיק, ואז אומר: "נעים מאוד, אהרון". המחשב יזהה אמירה זו כקוד הדורש ממנו לשמור את תמונתו של אהרון ליד השם "אהרון", ובפעם הבאה כבר יהיה אהרון בטוח שאני זוכר אותו (המחשב גם יוכל להזכיר לי היכן הייתי כאשר דיברתי עם אהרון בפעם האחרונה, ולהשמיע לי חלק משיחתנו באותה הזדמנות, אך זה נושא לכתבה אחרת).

גם כאן משתרע לפנינו מדרון חלקלק. יותר ויותר אנשים מעלים את ארכיון התמונות שלהם לאינטרנט. אם במקום כלשהו נמצאת התמונה שלי ליד שמי, זה מספיק כדי לחפש את פניי בכל תמונה אחרת המופיעה באתר כלשהו באינטרנט. יהיה אפשר גם לחפש אותי בהקלטות של שידורי כל מצלמת רשת (Web Cam) שהתמונות שהיא מצלמת נגישות לקהל הרחב. כיום זוכים רק סלבריטאים לזיהוי בכל אשר ילכו. האם בעתיד יהיה כל אחד מאיתנו אובייקט לחיפוש? האם כאשר יתכונן מישהו מאיתנו לפגישה חשובה, עסקית או חברתית, הוא יוכל למצוא באינטרנט לא רק היכן הופיע שמו של האדם שאיתו נפגש, אלא גם היכן היה בימים האחרונים (או גם לפני חמש שנים), היכן הוא נוהג לבלות, לאילו חנויות הוא נכנס וכו'?

מה יעשו פושעים בעזרת טכנולוגיה כזו? הם יוכלו לדעת מתי אתה רחוק ולכן אפשר לפרוץ לביתך, הם יוכלו להטריד ולהפחיד, הם יחפשו הזדמנויות לסחיטה, וודאי יחשבו על רעיונות יצירתיים נוספים. תסריטים מעוררי-חשש אלו אינם אפשריים עדיין, אך בקצב ההתקדמות הנוכחי, אין סיבה שלא נראה אותם מתממשים אפילו לפני סוף העשור.

הטכנולוגיה נותנת, הטכנולוגיה לוקחת

רלף גרוס (Gross), ממכון הרובוטיקה של אוניברסיטת קרנגי-מלון (Carnegie Mellon), מכיר היטב את שני הצדדים של הסוגיה. בין השאר, הוא עוזר למכוני רישוי בארצות-הברית לחפש אנשים שרישיון הנהיגה שלהם נשלל במדינה אחת, ואז ביקשו רישיון חדש במדינה אחרת, או תחת שם אחר באותה מדינה. מתברר כי על-ידי השוואת התמונות של אנשים שונים-לכאורה נתפסו כמה עבריינים כאלו, וייתכן שכך נמנעו כמה תאונות קטלניות.

מצד אחר, גרוס עובד עם מעבדת פרטיות המידע (Data Privacy Lab) של קרנגי-מלון כדי למצוא דרכים לעצור את ההחלקה במדרון. הגישה שהם מציעים מתבססת על עובדה מעניינת: אף שהן בני-אדם והן מחשבים מגיעים ליכולת גבוהה בזיהוי פנים, הם מבצעים אותה מטלה בדרכים שונות לחלוטין. למעשה, איננו יכולים לומר הרבה על השיטה האנושית או על השיטה הממוחשבת: לגבי בני-אדם, למרות ממצאים רבים על המקומות במוח שבהם מתבצע הזיהוי, ועל הגורמים למצבים פתולוגיים שבהם אובדת היכולת לזיהוי פנים (Prosopagnosia – מיוונית; פרוסופון פירושו "פנים" ואגנוסיה פירושה "אי-ידיעה"), עדיין איננו יודעים לתאר את עקרונות השיטה שבה פועל המוח (וראו: יפעת לוי ורפאל מלאך – "מקטן ועד גדול במוח האדם", "גליליאו" 49). מן הצד הטכנולוגי, אף שפורסמו בעבר כמה אלגוריתמים יעילים לזיהוי פנים, החברות שהשיגו את ההתקדמויות המרשימות שעליהן דווח ב-FRVT אינן חושפות את השיטות שבהן הגיעו לשיפורים אלו.

גם אם לא נוכל לבחון את השיטות השונות של האדם ושל המחשב, נוכל לראות כי במבחן התוצאה יש הבדלים: אנשים טובים יותר בזיהוי פנים שהם מכירים היטב (כלומר שכבר ראינו אותם בצורות שונות, בתאורות שונות וכו'), ובזיהוי פנים הנצפים מזוויות לא-שגרתיות (מהצד או מלמטה,



דמויות שונות. בסרט זה אפשר לראות אפילו הבעות פנים ולהבין את המאורעות, אך המשתתפים יוכלו להיות מזוהים רק לאחר קבלת צו שופט. תוך כדי הפיתוח התברר כי אפשר להשתמש בשיטה דומה גם כדי ליצור תמונות שבני-אדם מזוהים טוב יותר מאשר המחשב (אם כי עדיין הזיהוי לא יהיה וודאי) – נסו את ההדגמה (קישור בסוף הכתבה) כדי לראות כמה פעמים הניחוש שלכם מוצלח יותר משל התוכנה המתחרה בכם.

מחקר זה הוא חלק מתחום מחקר רחב וחשוב, הבודק שני צדדים של אותה שאלה: מצד אחד, המחקר מראה את היכולת של תוכנה לפרוץ את מעטה האנונימיות והפרטיות – אותו מעטה שאנו משלים את עצמנו אם אנו מאמינים כי הוא קיים וימשיך להתקיים בעתיד. מצד אחר, הוא מפתח כלים להסוואת זהות תוך פגיעה קטנה בכל האפשר בשימושיות של אותה טכנולוגיה העוזרת לנו ומגנה עלינו בשיטותינו ברשת, בתהליכים רפואיים, או אפילו כשאנו מהלכים ברחוב. לבינה המלאכותית יש חלק חשוב בשני צדדים אלו. ■

ישראל בנימיני עובד בחברת ClickSoftware בפיתוח שיטות אופטימיזציה מתקדמות.

למשל). מחשבים טובים יותר בזיהוי פנים שנראו רק פעם אחת וכאשר שתי העיניים נראות בבירור. שיטות המקשות על בני-אדם לזהות פנים אינן מקשות בהכרח על מחשבים: גרוס וחבריו לקבוצה מצאו כי הריבועים המסתירים את פני דמויות בטלוויזיה ובעיתונות עדיין מאפשרים לעתים קרובות לתוכנה לגלות את זהות המרואייין – אפילו ברמות טשטוש (גודל ריבוע) גבוהות במיוחד. לפעמים ריבועים אלו אפילו משפרים את יכולת הזיהוי של המחשב, מכיוון שהם מסירים פרטים שנוטים "לבלבל" את התוכנה! (בסוף הכתבה יש קישור למאמר על כך, ועל השיטה המוצגת בהמשך).

אם כך, האם אפשר למצוא תהליך הפוך – תהליך שימנע ממחשב לזהות את הפנים, אך עדיין לא יפגע בשימושיות של התמונה עבור בני-אדם? נראה כי קיימת שיטה כזו. חוקרים במעבדת פרטיות המידע, בראשות לטניה סוויני (Sweeney), פיתחו תהליך שבו משולבים פרטי הפנים של כמה אנשים שונים כדי ליצור תמונה חדשה, ותמונה זו מושלת על גבי התצלום או סרט הווידאו המקורי (קישור בסוף הכתבה). מטרתם היא להתמודד עם מצבים שבהם המשטרה, למשל, חושבת כי ייתכן שבסרטי מעקב מסוימים קיים מידע הקשור לחקירת פשע, אבל אין לה טיעונים מספיק משכנעים כדי ששופט יאשר גישה לאותם סרטים. הטיעונים יכולים להיות מוצגים רק לאחר שהמשטרה אכן תצפה בסרטים – מצב של "מלכוד 22".

הצוות של סוויני מציע שיטה להסוואת זהות: המשטרה תוכל לצפות בגרסה של הסרט שבה כל הפנים הוחלפו ב"פנים משולבים", כך שכל אחד מהם מורכב בצורה אחרת מכמה



קישורים

מבחן יצרני זיהוי הפנים:

<http://www.frvt.org>

גוגל רוכשת את נבן:

<http://googleblog.blogspot.com/2006/08/better-way-to-organize-photos.html>

מאמר על זיהוי דמויות "מפוקסלות" ועל הדרך הנכונה להסוואת זהות:

http://www.ri.cmu.edu/pub_files/pub4/gross_ralph_2006_2/gross_ralph_2006_2.pdf

הדגמה של שיטת הסוואת הזהות:

<http://privacy.cs.cmu.edu/query?htx=file://videox/bfind2.htx>