

# למידה עמוקה

# Deep Learning

מיקי אלעד  
מדעי המחשב - הטכניון



אז על מה מדובר ?

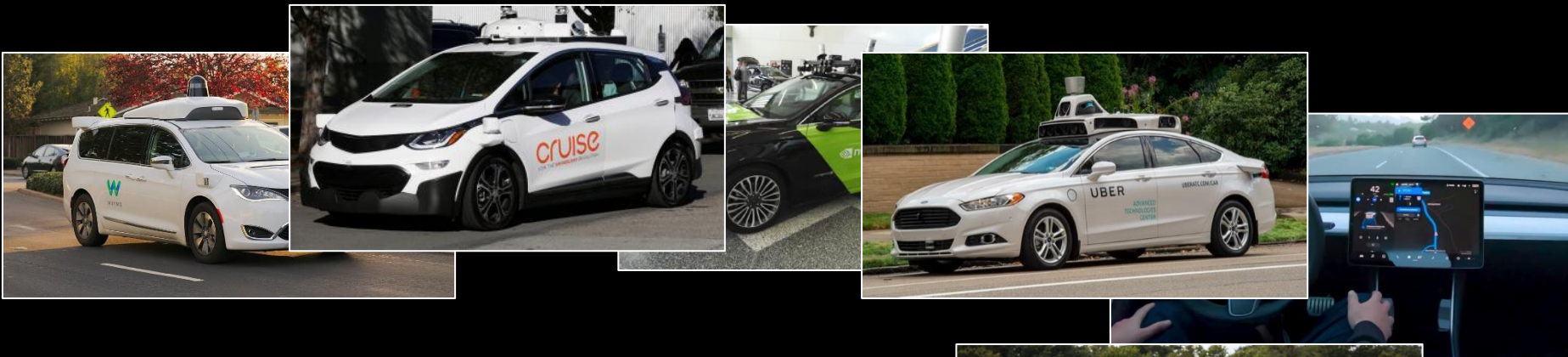
# מהפיכה !

זו אחת המהפכות הטכנולוגיות  
החשובות ביותר בעת האחרונה  
והיא צפויה להשפיע על כולנו ובמגוון אופנים

אז מה זה "למידה עמוקה" ?

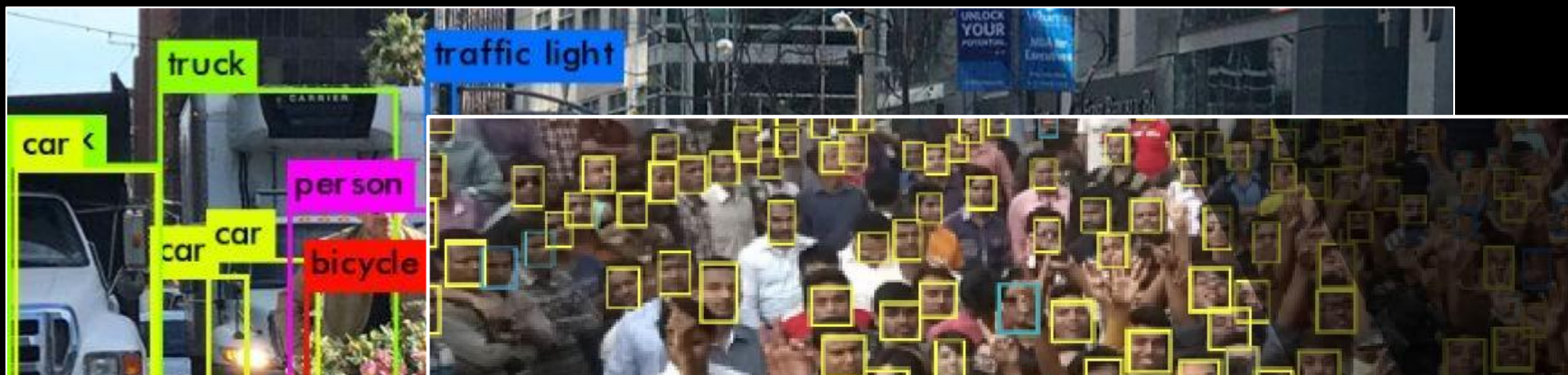
נתחיל ע"י הצגת כמה דוגמאות  
לחדירה של תחום זה והשפעתו

# נהיגה אוטונומית

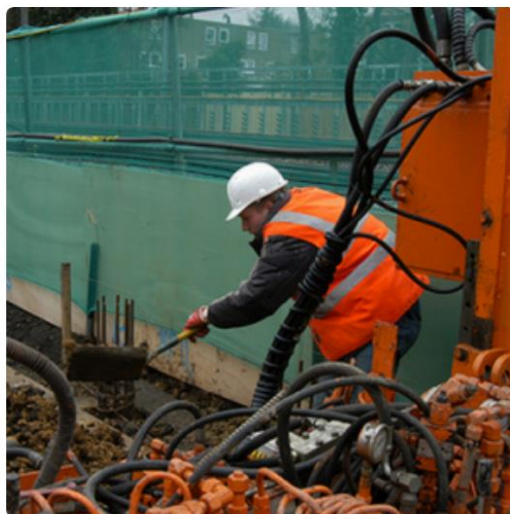


תאונות דרכים  
עמידת שוקקים  
מציאות מנייה  
בעלות על שוק

# זיהוי עצמים בתמונות



"man in black shirt is playing guitar."



"construction worker in orange safety vest is working on road."



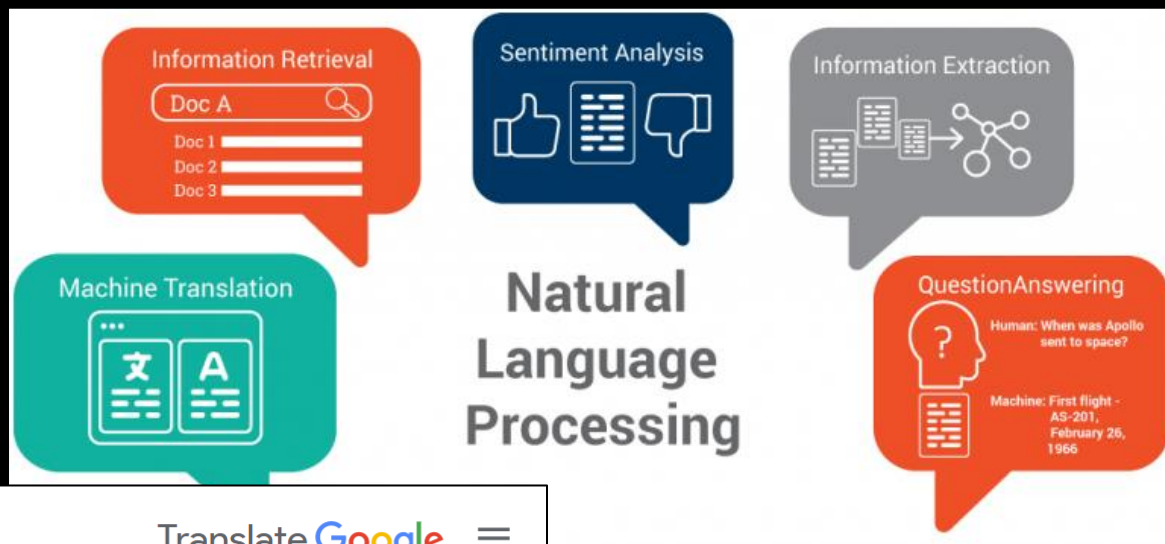
"two young girls are playing with lego toy."



# ניתוח טקסט ותרגום



תרגום סימולטני



Translate Google

Spanish Arabic **Hebrew** Detect Language ⇌ English Spanish Arabic Translate

עברית עברית

Translate your text here

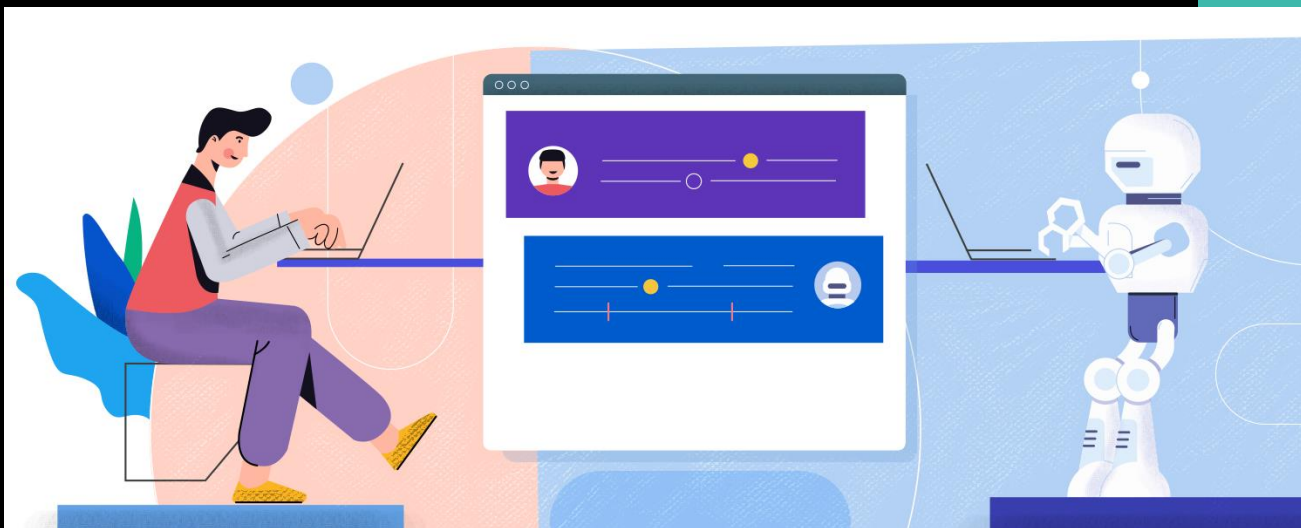
שלום לכולם, היום נדבר על מהפיכת  
הלמידה העמוקה והשפעתה על חיינו

Hello everyone, today we will talk  
about the profound learning  
revolution and its impact on our  
lives. 🔊

Translations by Microsoft® Translator

# ניהול שיחה אוטומטי

- האם מחשב יוכל לנהל שיחה (ע"י טקסט) עם אדם מבלי להיחשף ככזה? זהו **מבחן טיורינג** לאינטליגנציה
- כיום, בוטים עושים זאת, והם חלק מחיינו בשירות לקוחות, בשירותי רפואה, שירותי בנקאות, ועוד ...

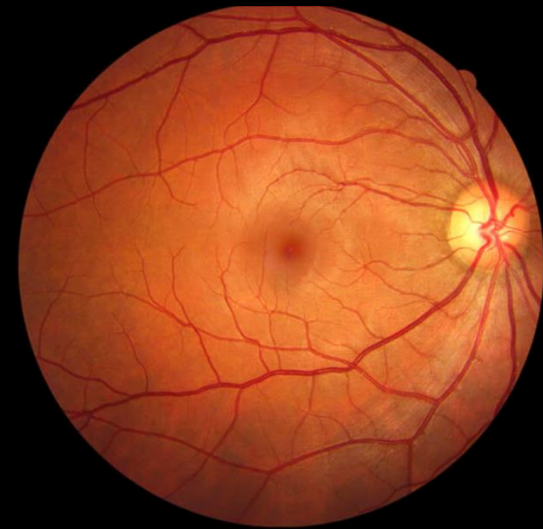


# אבחון מחלות

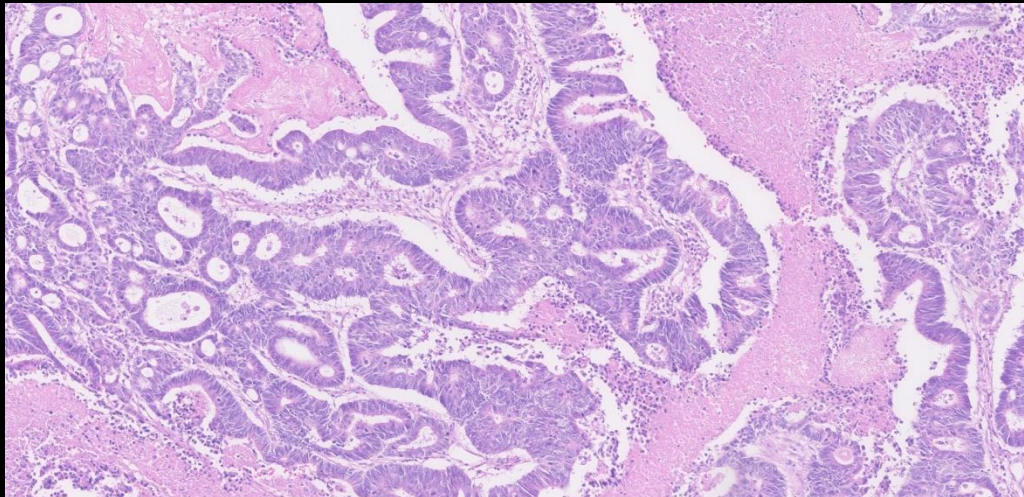


סטנפורד:  
אבחון נגעים  
בעור כסרטניים

Google-Brain: ניבוי  
סיכונים לאירוע לבבי או  
מוחי מתמונת הרשתית



BIDMC:  
זיהוי גידולים  
ברקמות פתולוגיה





# יצירת אומנות



Turner

Van Gogh

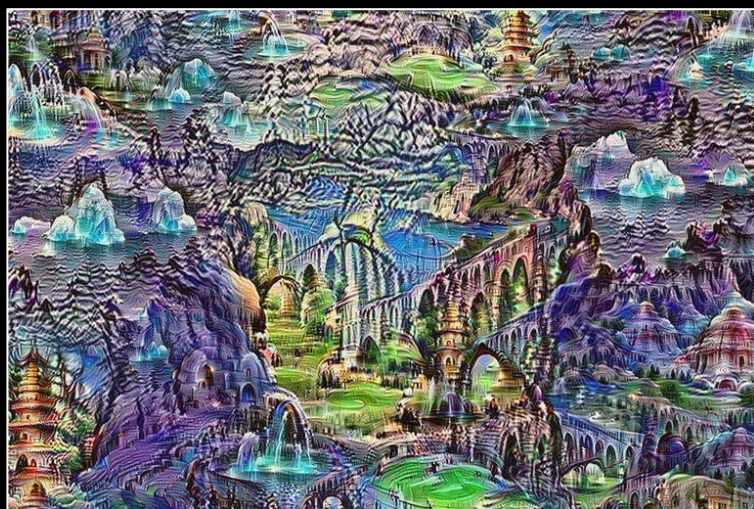
Munch

# יצירת יש-מאין



thispersondoesnotexist.com

תוכנה  
מבוססת  
למידה עמוקה  
יכולה לייצר  
מוזיקה,  
תמונות, טקסט  
או וידאו

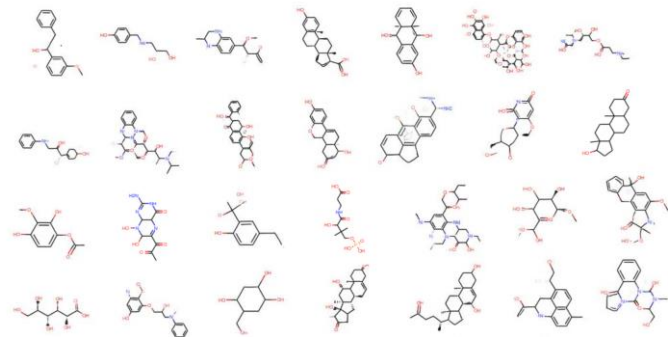


# ומה עוד ?

## Molecule synthesis using AI

Step by step guide to use GAN in search for new drugs and materials

Neeraj jain Sep 25, 2019 · 6 min read



07/02/2021 Mathematics

Using AI and computer automation, Technion researchers have developed a "conjecture generator" that creates mathematical conjectures, which are considered to be the starting point for developing mathematical theorems.

Using AI and computer automation, Technion researchers have developed a "conjecture generator" that creates mathematical conjectures, which are considered to be the starting point for developing mathematical theorems. They have already used it to generate a number of previously unknown formulas. The study, which was published in the journal Nature, was carried out by undergraduates from different faculties under the tutelage of Assistant Professor Ido Kaminer of the Andrew and Erna Viterbi Faculty of Electrical Engineering at the Technion.



רפואה – אבחנה אוטונומית, תכנון

תרופות, מחקר גנטי

איכות חיים – מערכות המלצה, סייע

וירטואלי, רבוטיקה

בטחון ואבטחה – צילום ומעקב, איסוף

וניתוח מודיעין

בנקאות – מסחר אלקטרוני,

גילוי רמאויות באשראי

תחבורה – ייעול תחבורה

ציבורית, נהיגה אוטונומית

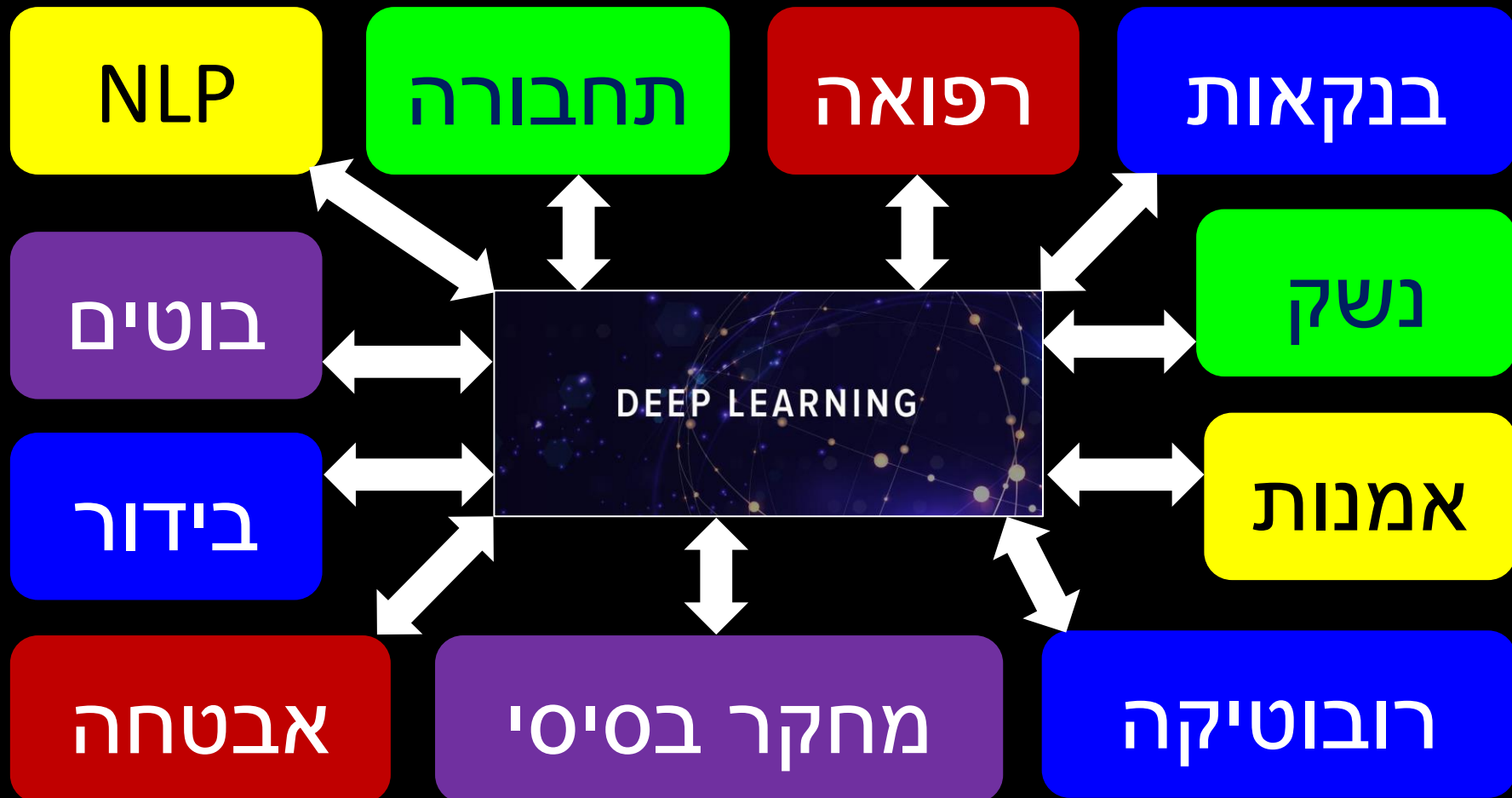
בידור – משחקי מחשב, סרטים

מחקר בסיסי – פיזיקה,

ביולוגיה, כימיה



# אכן מהפיכה !





# למידה עמוקה

כל אלה והרבה דוגמאות נוספות נשענים על הטכנולוגיה הקרויה

למידה עמוקה

אגב: התחום פופולרי וזוכה למגוון שמות אחרים כמו

מהפיכת ה-AI (Artificial Intelligence)

או

מהפיכת המידע

סיפורו של התחום

# למידה עמוקה ?

אז מה זה בדיוק?

מהיכן זה בא?

ויותר חשוב – לאן זה מתקדם?

אז הנה הדבר המפתיע:

זה היה כאן איתנו כבר בשנות השישים  
ואף אחד לא עמד על שעומד להתרחש

# חקר המוח האנושי

זה כבר עשרות שנים שחוקרים מייחלים  
להבין את דרך פעולתו של המוח האנושי  
ולהציע מערכת הנדסית שתחקה  
אותו – מכונה שתקבל מידע  
ותלמד ממנו על העולם עד  
לרמה שהיא תוכל להסיק  
מסקנות בעצמה

נשמע מסובך, נכון?

מסתבר שזה יכול להיות  
יותר פשוט ממה שחשבנו





# חקר המוח האנושי

החוקרים בזירה זו מתחלקים לשניים

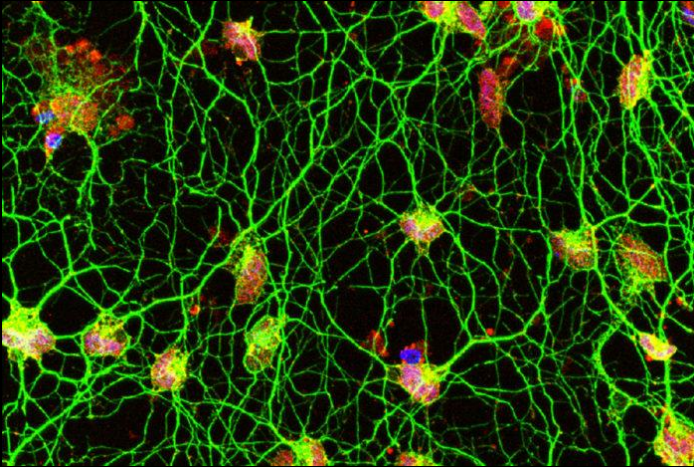


מדענים שלהם אין עניין  
בביולוגיה והם רוצים לפתור  
בעיות הנדסיות בדרכים יעילות

מדענים שבאים לעניין מתוך  
רצון לעסוק בחקר המוח  
ולהבין את דרכי פעולתו

הסיפור שלנו מתמקד בקבוצה השניה  
כי היא זו שמובילה את המהפכה העכשווית

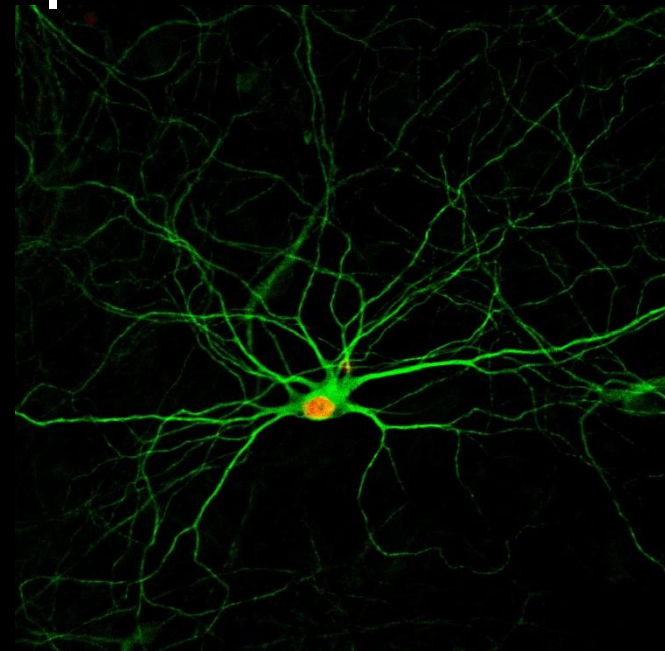
# הרעיון הבסיסי ?



המוח בנוי כרשת של נוירונים רבים ( $10^{11}$ ) שמקושרים ביניהם. כל נוירון הוא תא בודד וטיפש, אך לרשת כולה מתפתחת יכולת מדהימה של פתרון בעיות מורכבות

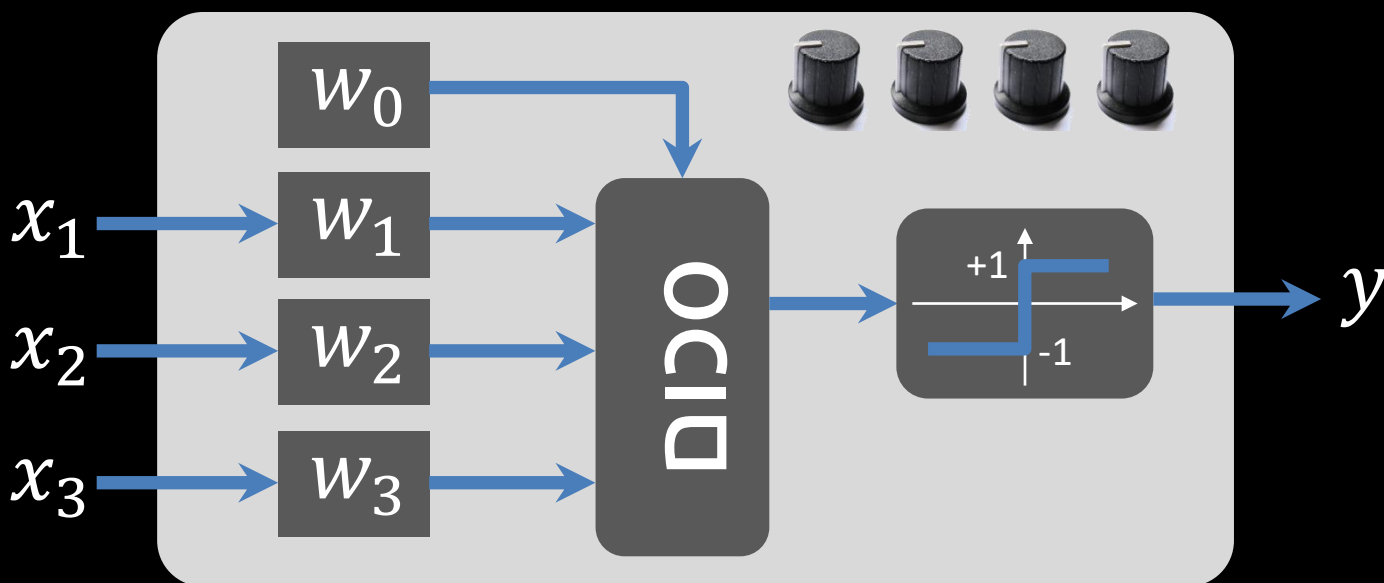
השאלה שמדענים עסקו בה היא:

האם נוכל להציע מקבילה  
דומה מעולם ההנדסה  
בעלת יכולות דומות?



# שנות השישים

בשנות השישים הוצע מענה לשאלה זו, כשאבן הבניין היא הפרספטרון (אנו נקרא לזה גם "נוירון")



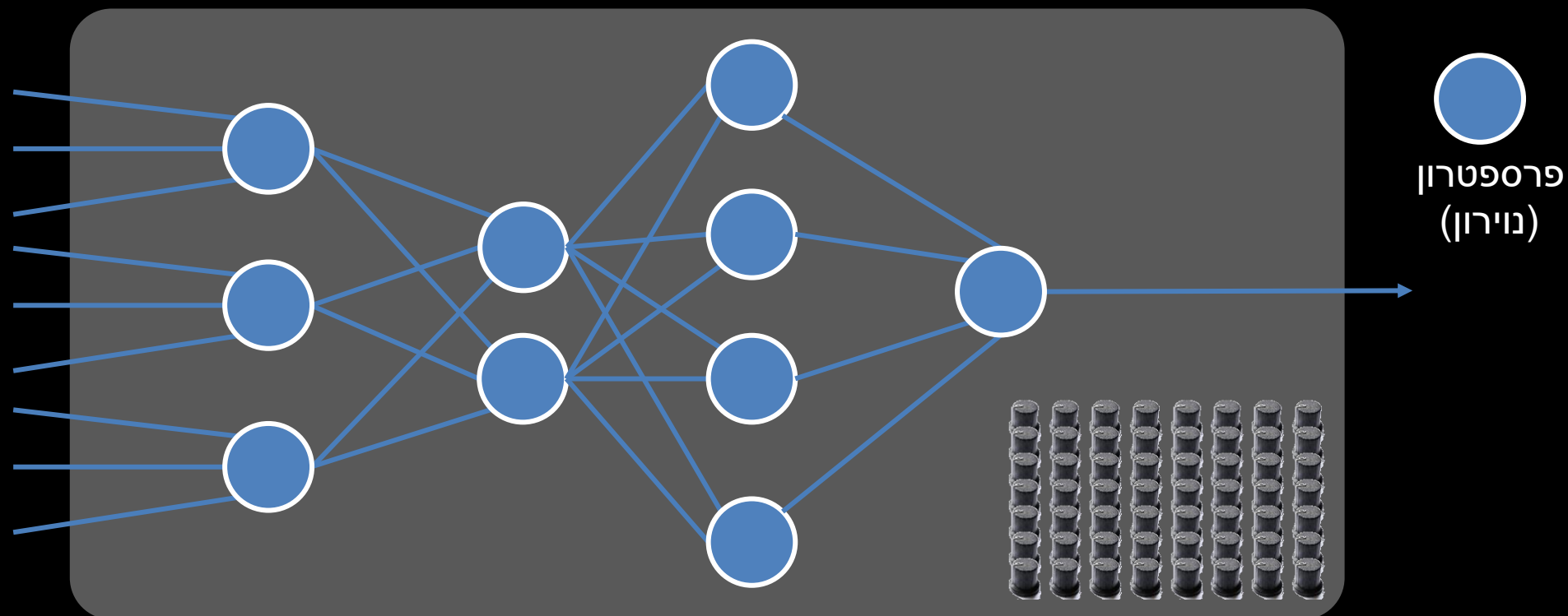
Frank Rosenblatt

$$y = \text{sign}\{w_0 + w_1x_1 + w_2x_2 + w_3x_3 + \dots\}$$

מכונה זו מבצעת ממוצע משוקלל של הכניסות והפעלת סף על התוצאה

# שנות השישים

המחשבה היתה שרשת של אלמנטים כאלו שמחוברים זה לזה  
(Feed-Forward) תוכל ליצור מערכת מורכבת ועשירה ביכולות



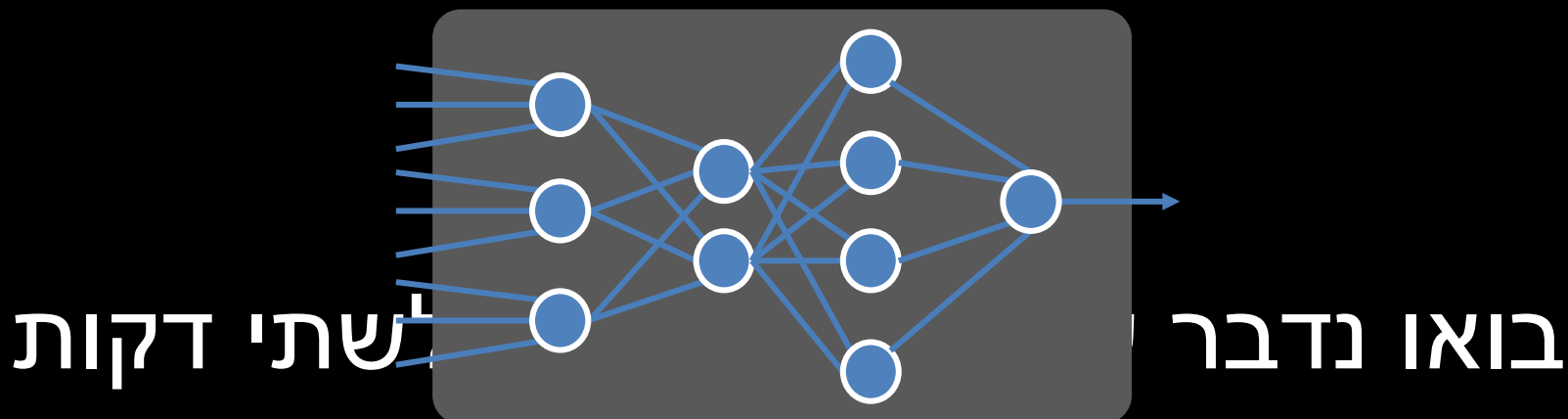
ניסו מגוון דברים בסיסיים עם רשתות כאלה, זה לא הצליח במיוחד  
ואף הציף בעיות שונות, ולכן עזבו זאת

# שנות השמונים

בשנות השמונים חזרו לנושא, חמושים בכלים עשירים וחזקים יותר (כמו אלגוריתם ה-backpropagation) והפעם ניסו

## ללמד את הרשתות

לבצע משימות סיווג והחלטה





# למידת מכונה ב-2 דקות

אנא הסתכלו בתמונה הבאה ואימרו

"האם זה **גבר** או **אשה**?"

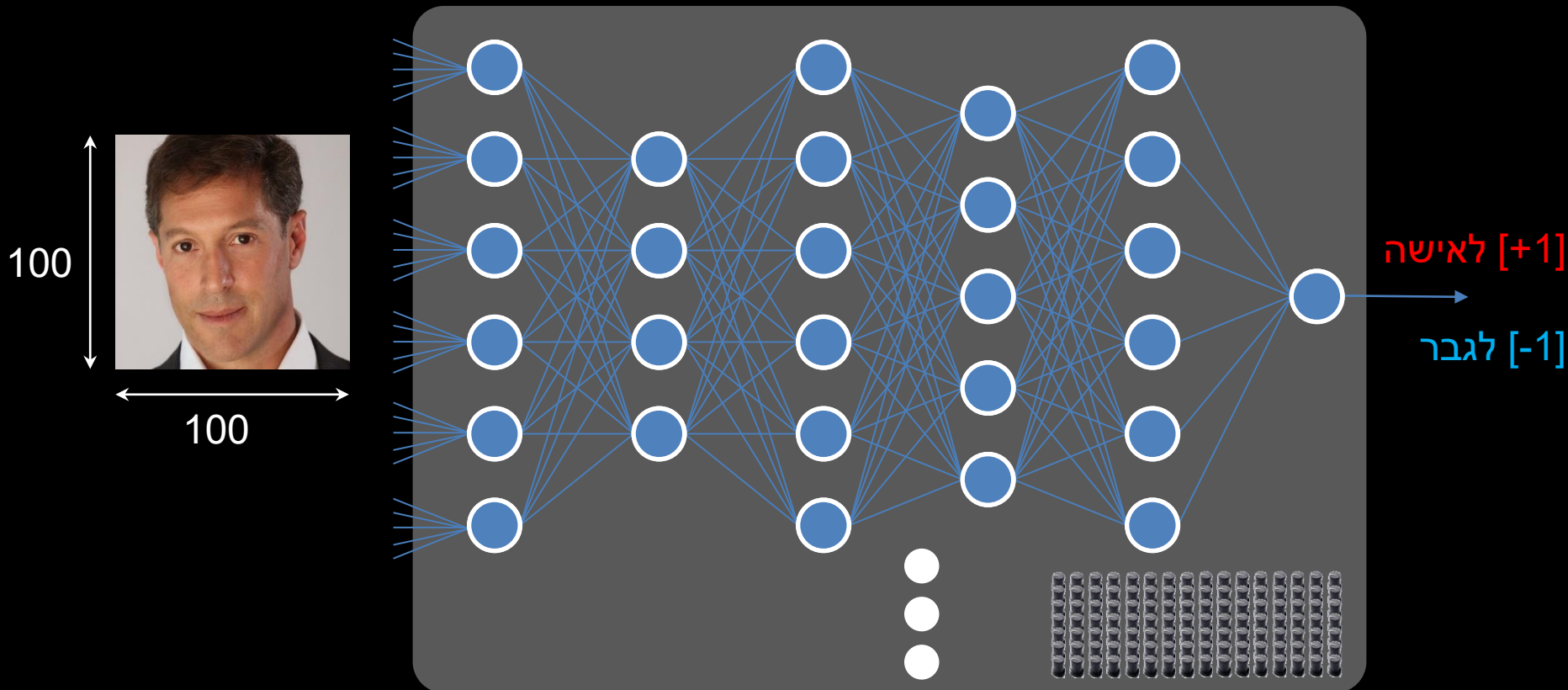


רוב הסיכוי שאקבל מכם תשובה נכונה.

האם נוכל ללמד רשת נוירונים לענות על שאלה כזו?

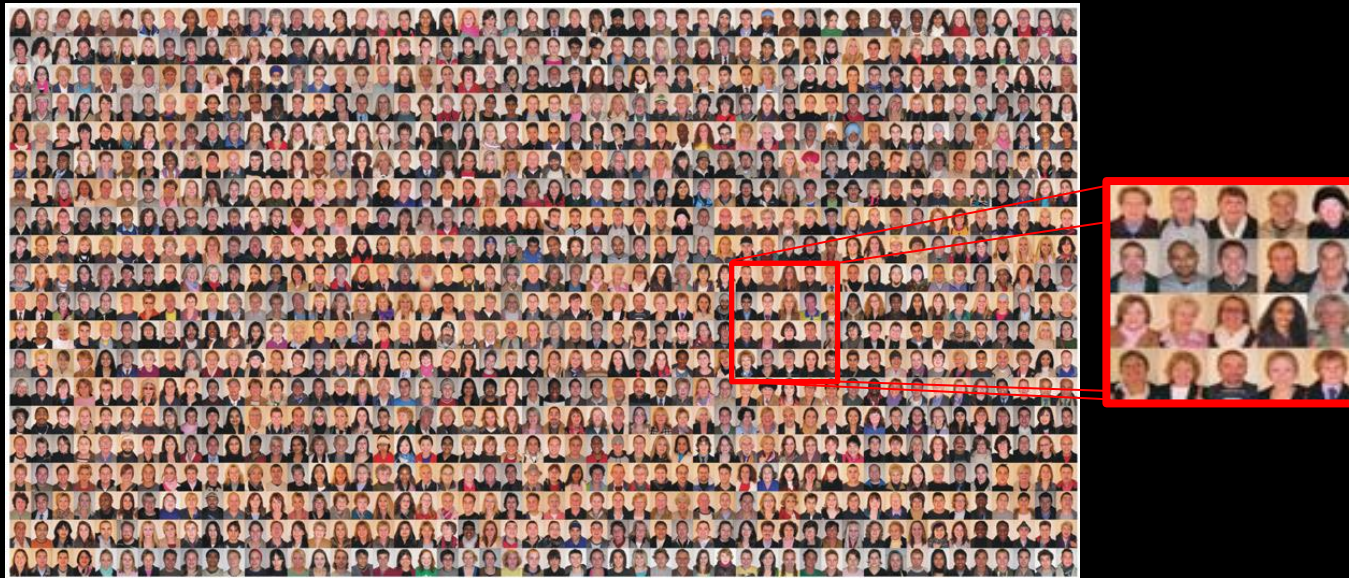
# למידת מכונה ב-2 דקות

המשימה שלנו היא לבנות רשת שבכניסתה תוכנס תמונה (שאיננה אלא אוסף מספרים), אוסף הנוירונים יעשה כל מיני חישובים ולבסוף יוציא במוצא  $[+1]$  לאישה ו-  $[-1]$  לגבר



# למידת מכונה ב-2 דקות

**אימון:** נאסוף תמונות פנים רבות של גברים ונשים (עם סיווגם)  
ו-"נכוון את פרמטרי הרשת" כך שיזוהו נכון



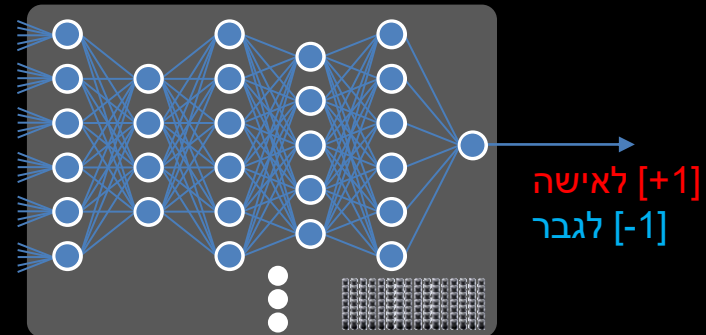
**בדיקה:** כשהאימון הסתיים, נוכל להזין תמונות חדשות לרשת  
והיא תנפיק ניבוי של המצולם – גבר או אשה. אם הדיוק  
על תמונות אלה גבוה נאמר כי יש "יכולת הכללה טובה"

# למידת מכונה ב-2 דקות

בואו נדבר מעט יותר על "כוונן פרמטרי הרשת"

זהו תהליך הלמידה שבו משתמשים בדוגמאות  
על מנת לאמן את רשת הנוירונים להגשים את משימתה

1. נניח שלרשת שלנו יש 50,000 כפתורים והמשימה היא לכוונן להשגת דיוק מירבי
2. ניקח את הרשת שבידינו ונזין לה את כל הדוגמאות (נניח 10,000) ונמדוד את דיוקה עבורן – נניח שהתקבל 15% כלומר 1,500 פרצופים סווגו נכון



# למידת מכונה ב-2 דקות

בואו נדבר מעט יותר על "כוונן פרמטרי הרשת"

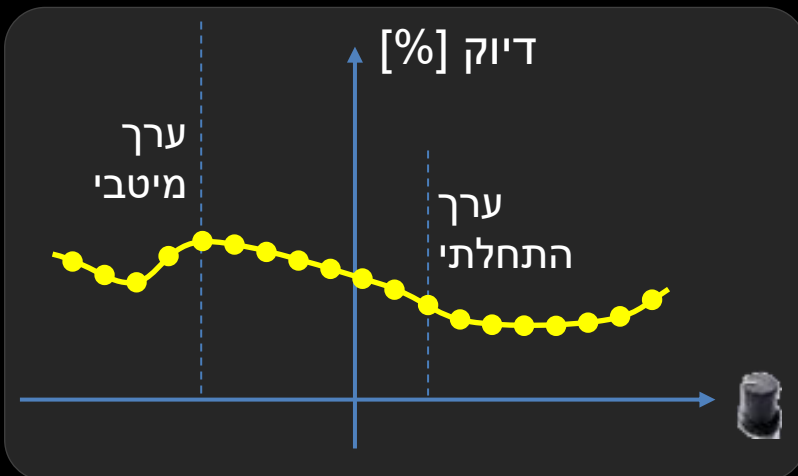
זהו תהליך הלמידה שבו משתמשים בדוגמאות על מנת לאמן את רשת הנוירונים להגשים את משימתה

3. נבחר את כפתור מס' 1 ברשת ו-"נשחק איתו", כלומר נשנה אותו ונמדוד את הדיוק בתלות בערכו, ונקבל את הגרף הבא:

4. נקבע את כפתור זה על ערך ש"סוחט" את הדיוק המירבי.

5. נחזור על כל תהליך זה שוב ושוב ... ושוב לכל כפתור ובמחזורים שלמים של סריקה

6. בהדרגה, ביצועי הרשת ישתבחו עוד ועוד עד לנגיע לנקודה שבה כל שינוי של כל כפתור רק גורע – זה הזמן לעצור ולהצהיר על סיום האימון





# למידת מכונה ב-2 דקות

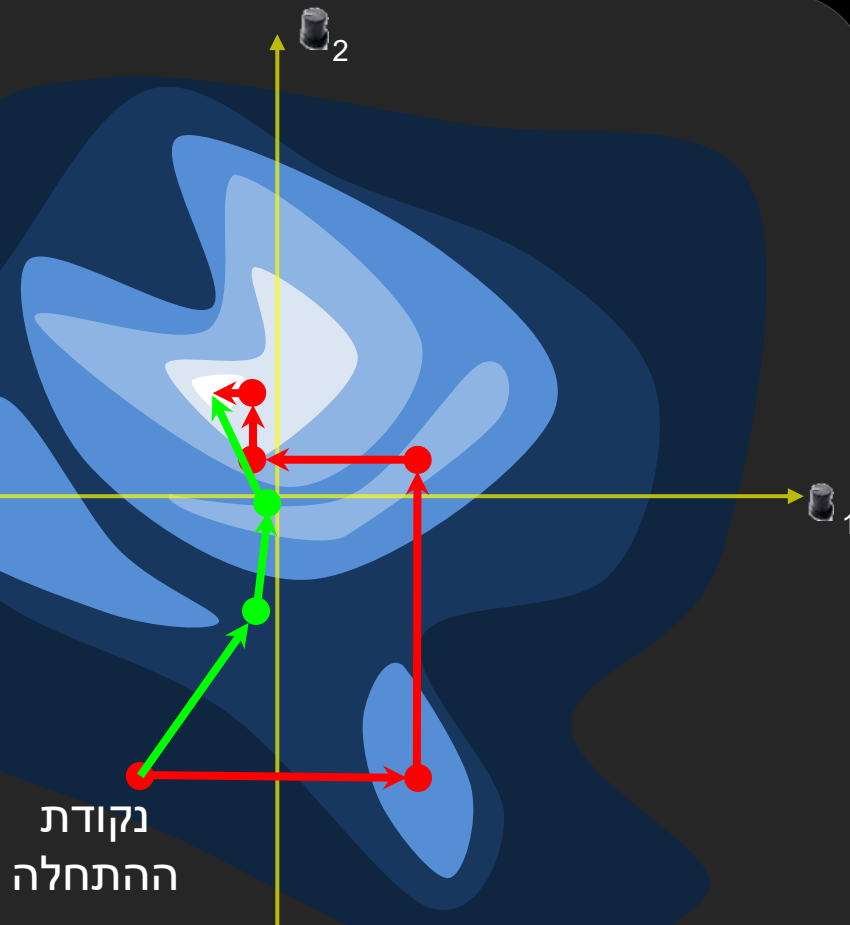
□ מה עשינו בתהליך אימון זה? חיפשנו תצורה של ערכי הכפתורים שתביא לדיוק מקסימלי, כלומר פתרנו בעיית אופטימיזציה

□ התהליך שתואר זה עתה מוכר בשם אלגוריתם ה- **Coordinate Ascent** והנה המחשה שלו בדו-מימד

□ לא ככה מאמנים באמת. אז מה עושים בפועל?

□ משתמשים באלגוריתם אחר שנקרא **Steepest Ascent** - שיטה שדורשת חישוב נגזרת

□ ה- **Backpropagation** שהוזכר קודם נועד לאפשר חישוב נגזרת על רשת עמוקה





# חזרה אל שנות השמונים

בשנות השמונים כל זה היה ידוע, ולימוד באופן זה רשתות  
נוירונים למגוון בעיות סיווג. הבעיות שבהן נתקלו כללו

– ביצועים בינוניים

– תחרות קשה מצד שיטות למידה אחרות (SVM)

– חוסר "אלגנטיות" מתמטית

– צריכת חישובים גדולה מידי

ושוב נכנס התחום לתרדמה, ואף סבל מ...

**רתיעה אגרסיבית**

# נעבור לשנות האלפיים

הנושא חזר לשולחן בשל כמה חוקרים עקשניים



Yan LeCun



Geoffery Hinton



Yoshua Bengio

אלא שהפעם הניסויים הובילו לתוצאות מעולות  
למה? מה השתנה?

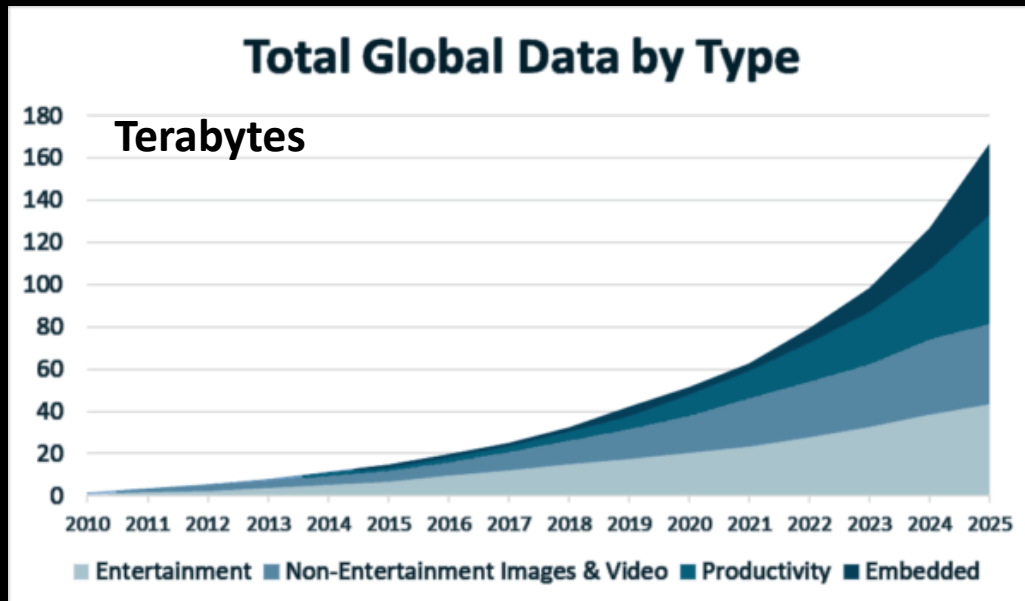
# אז מה השתנה ?

קרו שני דברים דרמטיים שאינם  
קשורים בהכרח ללמידה עמוקה:

1. **כמות המידע גדלה פלאים**, בשל האינטרנט, המצלמות

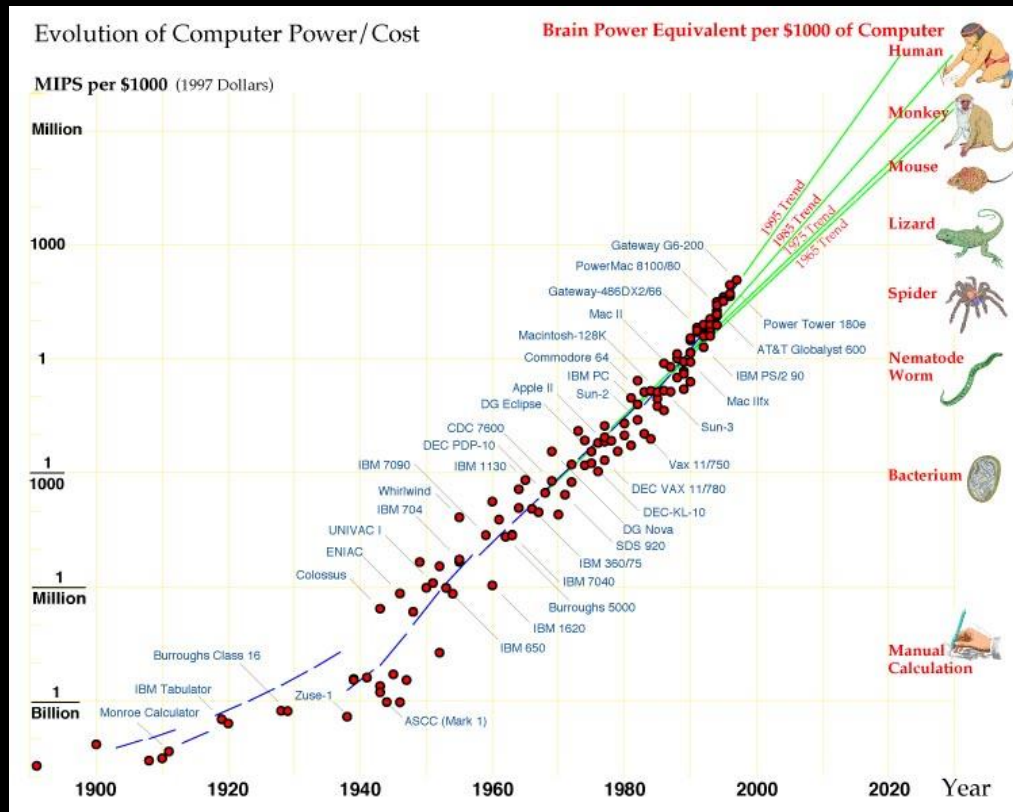
הדיגיטליות, ובעיקר  
בשל יכולת האגירה  
המשתפרת

בשנות ה-80 זיכרון  
מחשב נמדד ב-kB  
ואילו בשנות ה-2000  
זה הומר ב-GB  
(פי מיליון)



# אז מה השתנה ?

קרו שני דברים דרמטיים:



## 2. מחשבים השתפרו

מאוד ביכולות  
החישוב שלהם

מחשב קונוונציונלי בשנת  
2000 היה נחשב כמחשב-על  
בשנות ה-80!

תכנית מחשב שדרשה זמן  
ריצה של שנה שלמה יכולה  
להתבצע כיום בחצי דקה !!!

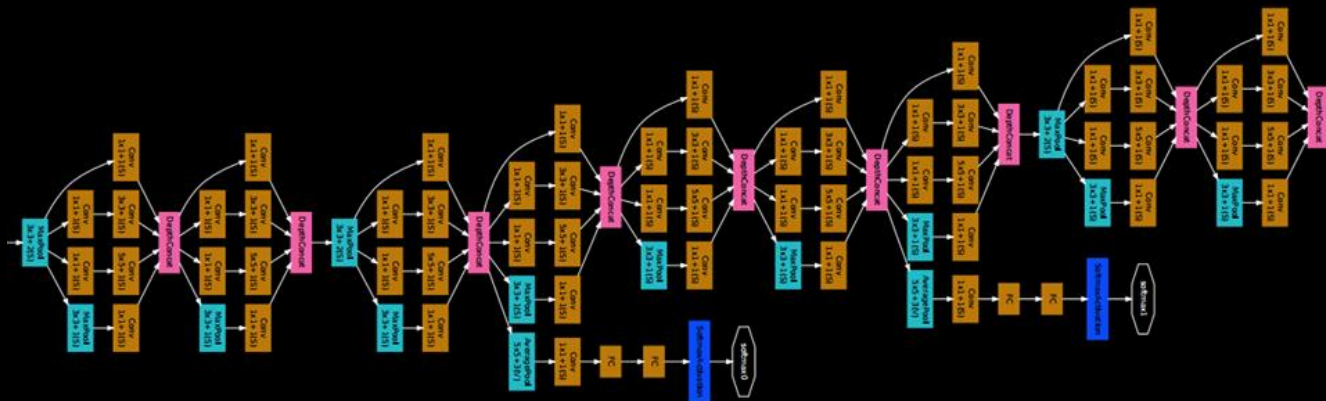
# מה קרה בשנות האלפיים

וכך, כשחזרו ואימנו רשתות לפי אותם עקרונות אבל

1. עם כמות מידע לאימון גדולה פי אלפים

2. עם שימוש במחשבים חזקים המהירים פי אלפים ויותר מקודמיהם, ו-

3. תוך שימוש ברשתות עם הרבה שכבות (ולכן השם "למידה עמוקה")



הרשתות הלומדות הגיעו לרמת ביצוע  
מעולה וחסרת תחרות



# נקודת המפנה

## ImageNet Classification with Deep Convolutional Neural Networks

**Alex Krizhevsky**  
University of Toronto  
kriz@cs.utoronto.ca

**Ilya Sutskever**  
University of Toronto  
ilya@cs.utoronto.ca

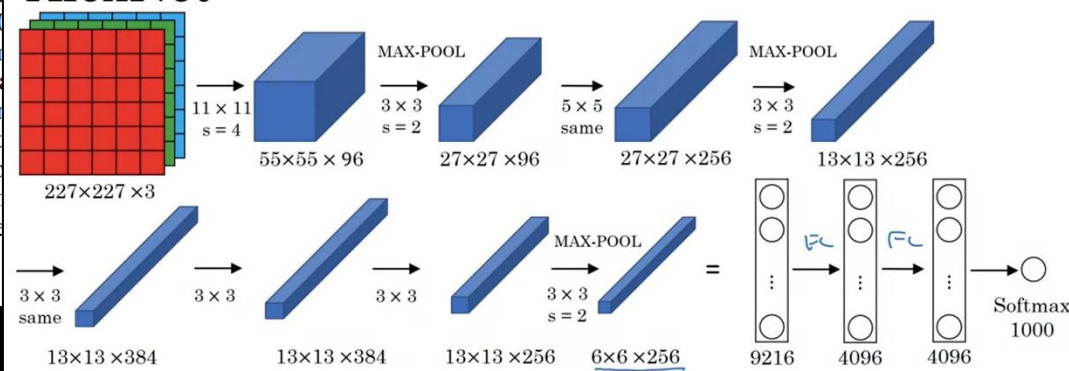
**Geoffrey E. Hinton**  
University of Toronto  
hinton@cs.utoronto.ca

### Abstract

We trained a large, deep convolutional neural network to classify the 1.2 million high-resolution images in the ImageNet LSVRC-2010 contest into the 1000 different classes. On the test data, we achieved top-1 and top-5 errors of 37.5% and 17.0% which is considerably better than the previous state-of-the-art. Our network consists of five convolutional layers, some of which are followed by max-over-time/pooling layers, and three fully-connected layers with a final 1000-way softmax. To make training faster, we used non-saturating neurons and a very efficient implementation of the convolution operation. To reduce overfitting in the training layers we employed a recently-developed regularization method called dropout that proved to be very effective. We also entered a variant of our network in the ILSVRC-2012 competition and achieved a winning top-5 test error rate of 26.2% compared to 26.2% achieved by the second-best entry.

הרגע המכונן בו כבר  
הובן שרשתות נירונים  
עמוקות מביאות  
מהפיכה הוא מאמר  
זה משנת 2012

### AlexNet



המצב כיום

# התעשייה חוגגת !

שפע של בעיות שנדמו כלא פתירות זכות לטיפול יעיל וקל  
ואף מראות הצלחה כבירה, ובלבד שייאסף מספיק  
מידע לאימון הרשת. זה נוגע במגוון תחומים שכבר הזכרנו

Google  
Tensor Processing Unit



אנקדוטה: Google

- Google Research → Google AI
- פיתוח תוכנת ה-TensorFlow
- פיתוח חומרת ה-TPU
- כמות עצומה של פרויקטים שבהם למידה עמוקה מהווה את עמוד השדרה המרכזי



TensorFlow

# התעשייה חוגגת !

מי עוד במשחק?



# התעשייה חוגגת!

מי עוד במשחק?

## ISRAEL'S ARTIFICIAL INTELLIGENCE STARTUPS

\$3.2B

### TECHNOLOGIES

\$310M

|                                |                                    |                          |                                  |                             |
|--------------------------------|------------------------------------|--------------------------|----------------------------------|-----------------------------|
| <b>Image/Video Recognition</b> | <b>Data Science &amp; Enablers</b> | <b>Speech &amp; Text</b> | <b>Chatbots &amp; Assistants</b> | <b>Security &amp; Other</b> |
|                                |                                    |                          |                                  |                             |

### SECTORS

\$100M

|                 |                 |                  |                               |                                   |
|-----------------|-----------------|------------------|-------------------------------|-----------------------------------|
| <b>Shopping</b> | <b>Agritech</b> | <b>Cleantech</b> | <b>Education/Legal/Sports</b> | <b>Social, Travel &amp; Other</b> |
|                 |                 |                  |                               |                                   |

### INDUSTRIAL

\$113M

### AUTOMOTIVE

\$35M

|                    |                     |                                |              |                   |                 |              |
|--------------------|---------------------|--------------------------------|--------------|-------------------|-----------------|--------------|
| <b>Maintenance</b> | <b>Optimization</b> | <b>Security &amp; Robotics</b> | <b>Other</b> | <b>Autonomous</b> | <b>Security</b> | <b>Other</b> |
|                    |                     |                                |              |                   |                 |              |

### ENTERPRISE

\$1.1B

|                 |                                  |                             |                            |
|-----------------|----------------------------------|-----------------------------|----------------------------|
| <b>Security</b> | <b>Internal Data &amp; Intel</b> | <b>Development &amp; IT</b> | <b>Support &amp; Other</b> |
|                 |                                  |                             |                            |

### HEALTHCARE

\$240M

|                |                    |                |                                 |                  |                              |
|----------------|--------------------|----------------|---------------------------------|------------------|------------------------------|
| <b>Imaging</b> | <b>Diagnostics</b> | <b>Monitor</b> | <b>Analytics &amp; Insights</b> | <b>Assistant</b> | <b>Discovery &amp; Other</b> |
|                |                    |                |                                 |                  |                              |

### FINTECH

\$261M

|                                |                              |                              |                             |                       |
|--------------------------------|------------------------------|------------------------------|-----------------------------|-----------------------|
| <b>Trading &amp; Investing</b> | <b>Market &amp; Insights</b> | <b>Insurance &amp; Fraud</b> | <b>Personal &amp; Other</b> | <b>Sales &amp; HR</b> |
|                                |                              |                              |                             |                       |

### MARKETING

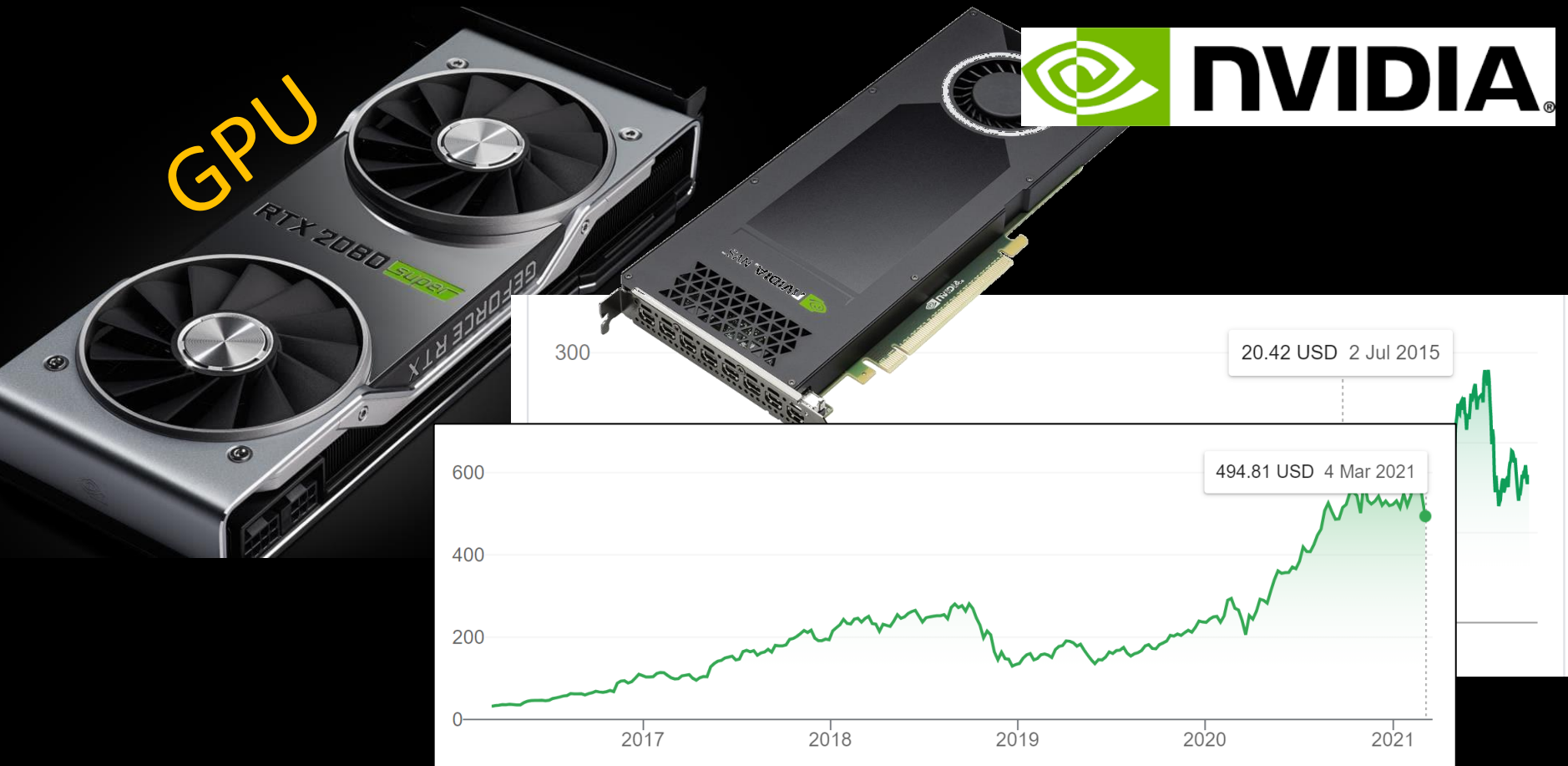
\$510M

|                           |                             |                                 |                         |                |                 |                |                             |                                    |
|---------------------------|-----------------------------|---------------------------------|-------------------------|----------------|-----------------|----------------|-----------------------------|------------------------------------|
| <b>Advanced Targeting</b> | <b>Predictive Analytics</b> | <b>Analytics &amp; Insights</b> | <b>Mobile Marketing</b> | <b>Bidding</b> | <b>Platform</b> | <b>Content</b> | <b>Customer Interaction</b> | <b>Lead Generation &amp; Other</b> |
|                           |                             |                                 |                         |                |                 |                |                             |                                    |



# מה קרה בשנות האלפיים

אנקדוטה: חישוב מקבילי וחברת NVIDIA





# מה קרה בשנות האלפיים

NVIDIA CEO Jensen Huang to Host AI Pioneers Yoshua Bengio, Geoffrey Hinton and Yann LeCun, and Others, at GTC21

Online Conference to Feature Jensen Huang Keynote and 1,300 Talks from Leaders in Data Center, Networking, Graphics and Autonomous Vehicles

Tuesday, March 9, 2021



ידיעה מהימים  
האחרונים

# תקופה מאתגרת למדע

## האקדמיה בוכה!

לי ולחבריי המדענים התקופה ... מאתגרת וקשה

אחת הבעיות עם תחום הלמידה העמוקה הוא ...

**רדידותו המחרידה**

המסר העיקרי: לא נדרש להבין בעיה על מנת לפותרה

רדידות זו גם מתאפיינת בכך שהעבודה בתחום זה

**היא נסיונאית באופיה, וללא הבנה תיאורית מעמיקה**

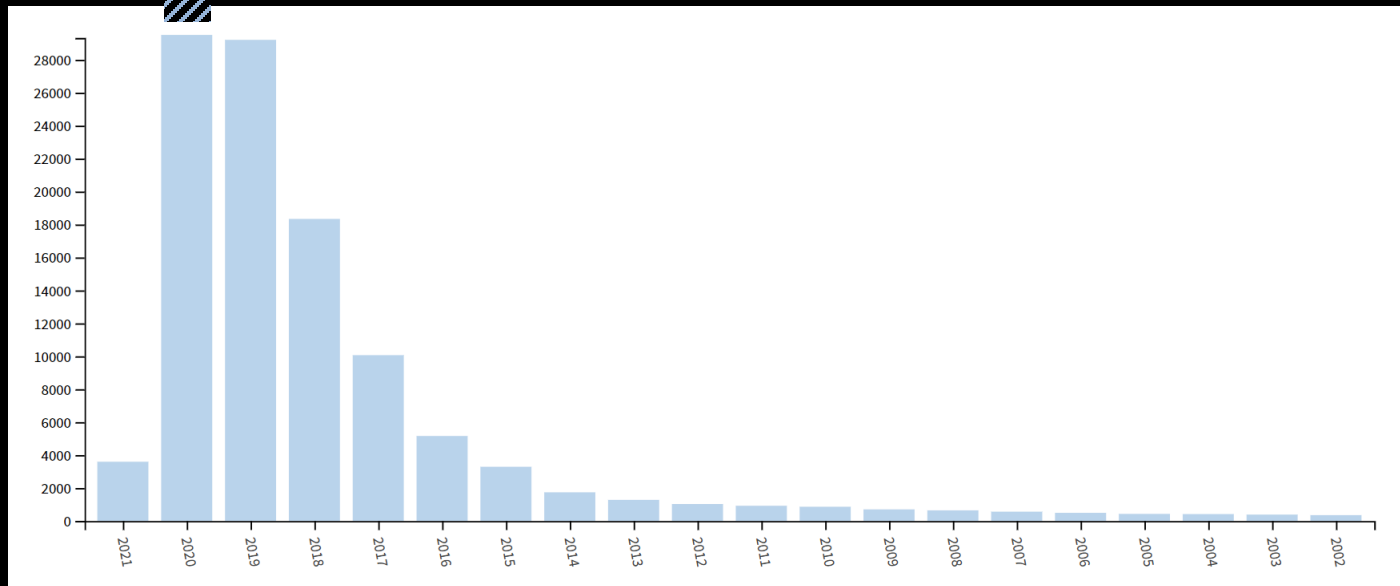
ואם כל זה לא מספיק, רבים מהמדענים

**סובלים מאבדן דרך, וחוסר הבנה למהותו של מחקר**

# תקופה מאתגרת למדע

אל תטעו! האקדמיה מסתערת בשצף-קצף על תחום זה ויש הצפה של עבודות מחקר בתחום של למידה עמוקה

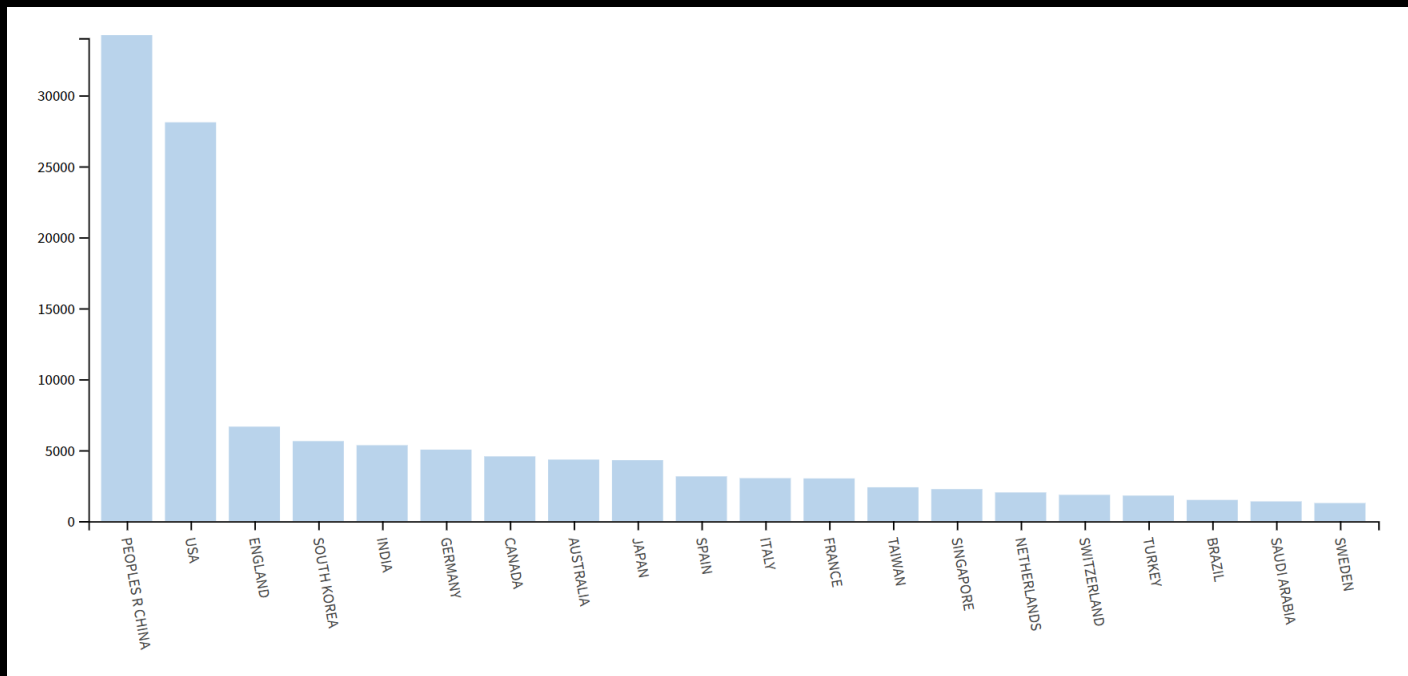
מספר המאמרים שפורסמו כל שנה בתחום



# תקופה מאתגרת למדע

אל תטעו! האקדמיה מסתערת בשצף-קצף על תחום זה ויש הצפה של עבודות מחקר בתחום של למידה עמוקה

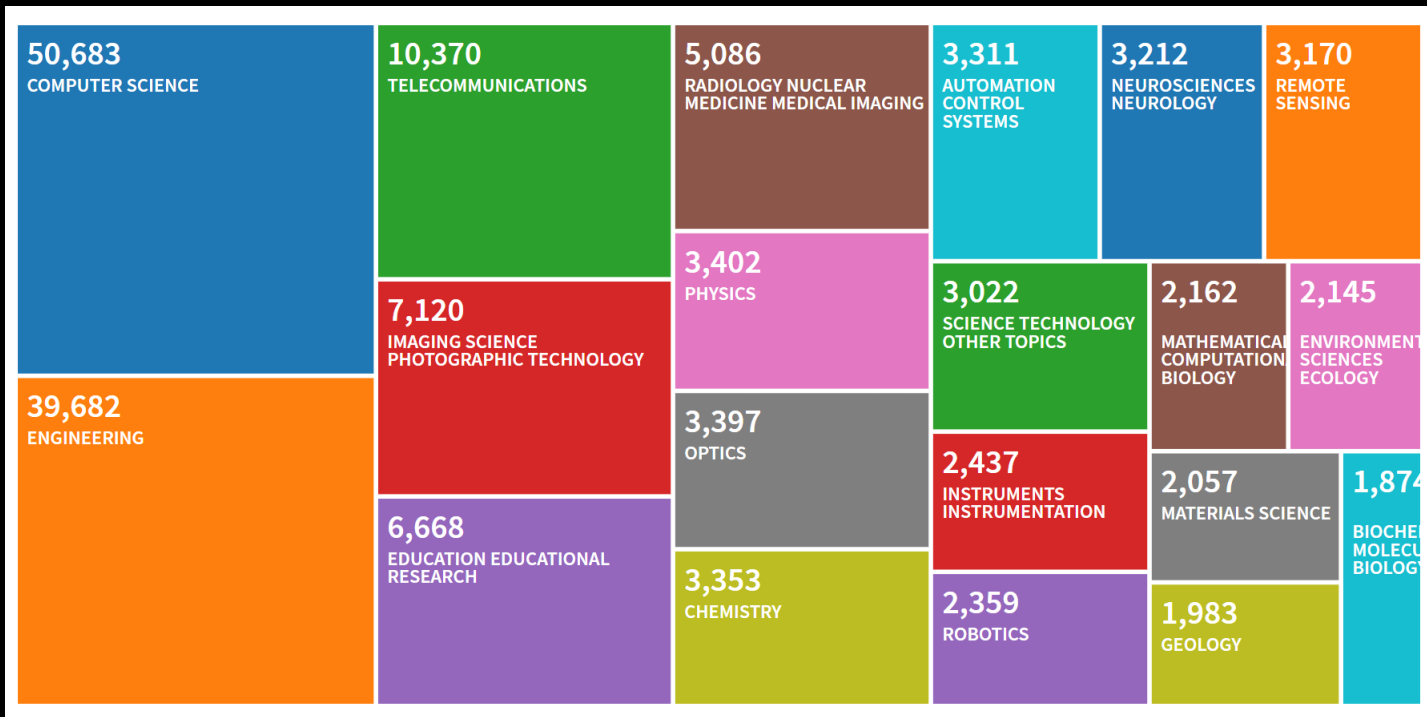
## מקור המאמרים שפורסמו בתחום



# תקופה מאתגרת למדע

אל תטעו! האקדמיה מסתערת בשצף-קצף על תחום זה ויש הצפה של עבודות מחקר בתחום של למידה עמוקה

תתי התחומים שמאמרים אלו עוסקים בהם



# תקופה מאתגרת למדע

אל תטעו! האקדמיה מסתערת בשצף-קצף על תחום זה ויש  
הצפה של עבודות מחקר בתחום של למידה עמוקה

**אבל ...**

רבים ממאמרים אלו הם בדרך כלל

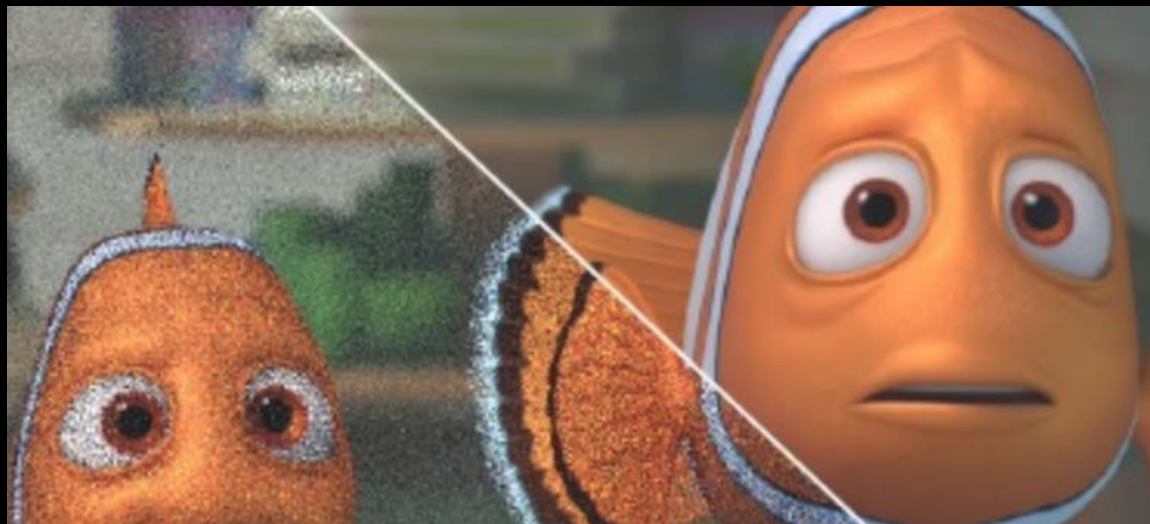
טכניים ואפרפריים בעיקרם

פותרים בעיות הנדסיות עם שימוש בהמון היוריסטיקות ובחירות שרירותיות  
עמוסי ניסויים (רוב המאמר יתאר ניסויים, ואילו החלק "החכם" יהיה קצר ודל)

ובסה"כ מאמרים אלו לרוב משאירים אותנו באפילה  
באשר לסיבות ההצלחה שלהם



# הצצה מקרוב – מחקר?

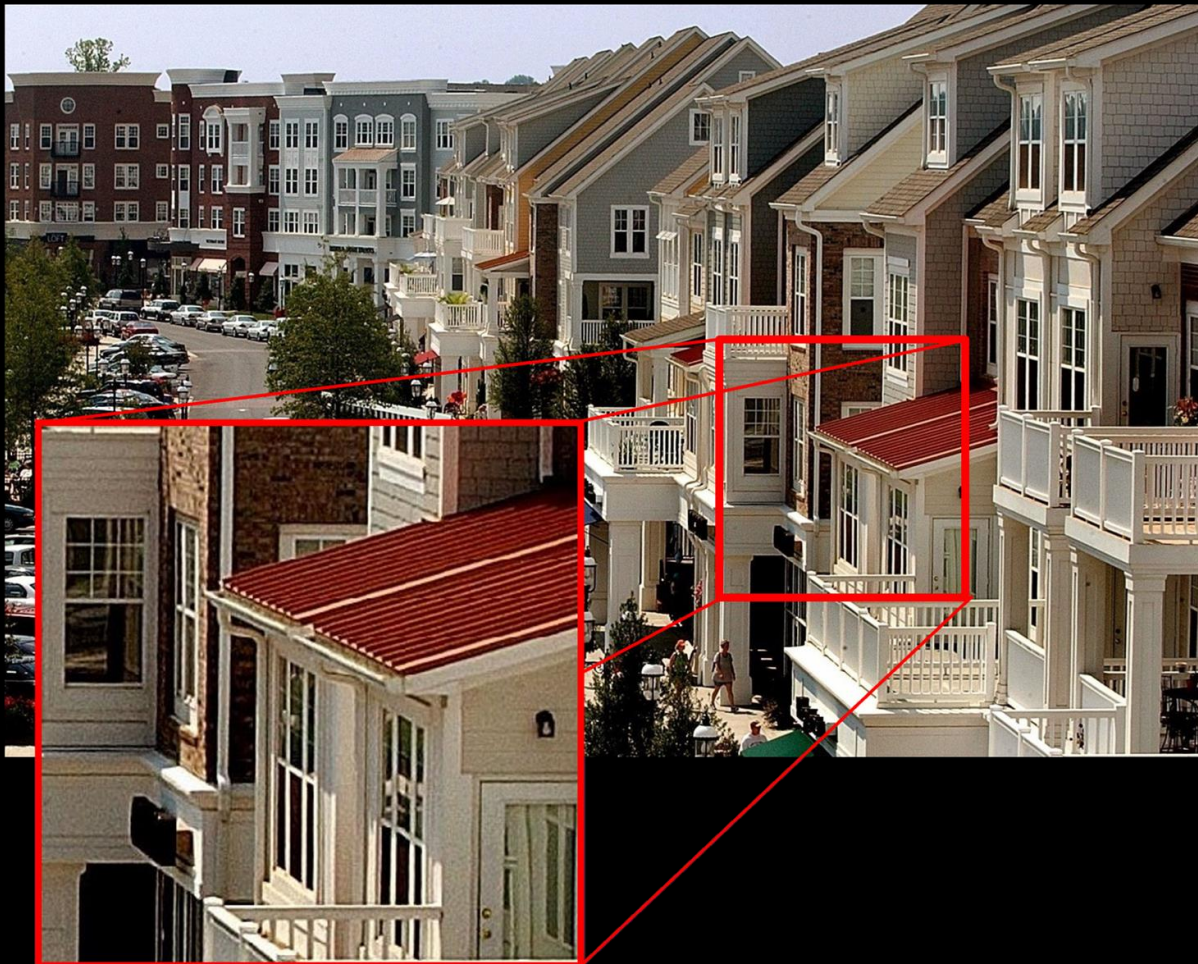


תנו לי להדגים לכם את  
הבעיות בתחום דרך  
נושא שעליו עבדתי ואני  
עדיין עובד שנים:  
**ניקוי רעש בתמונות**

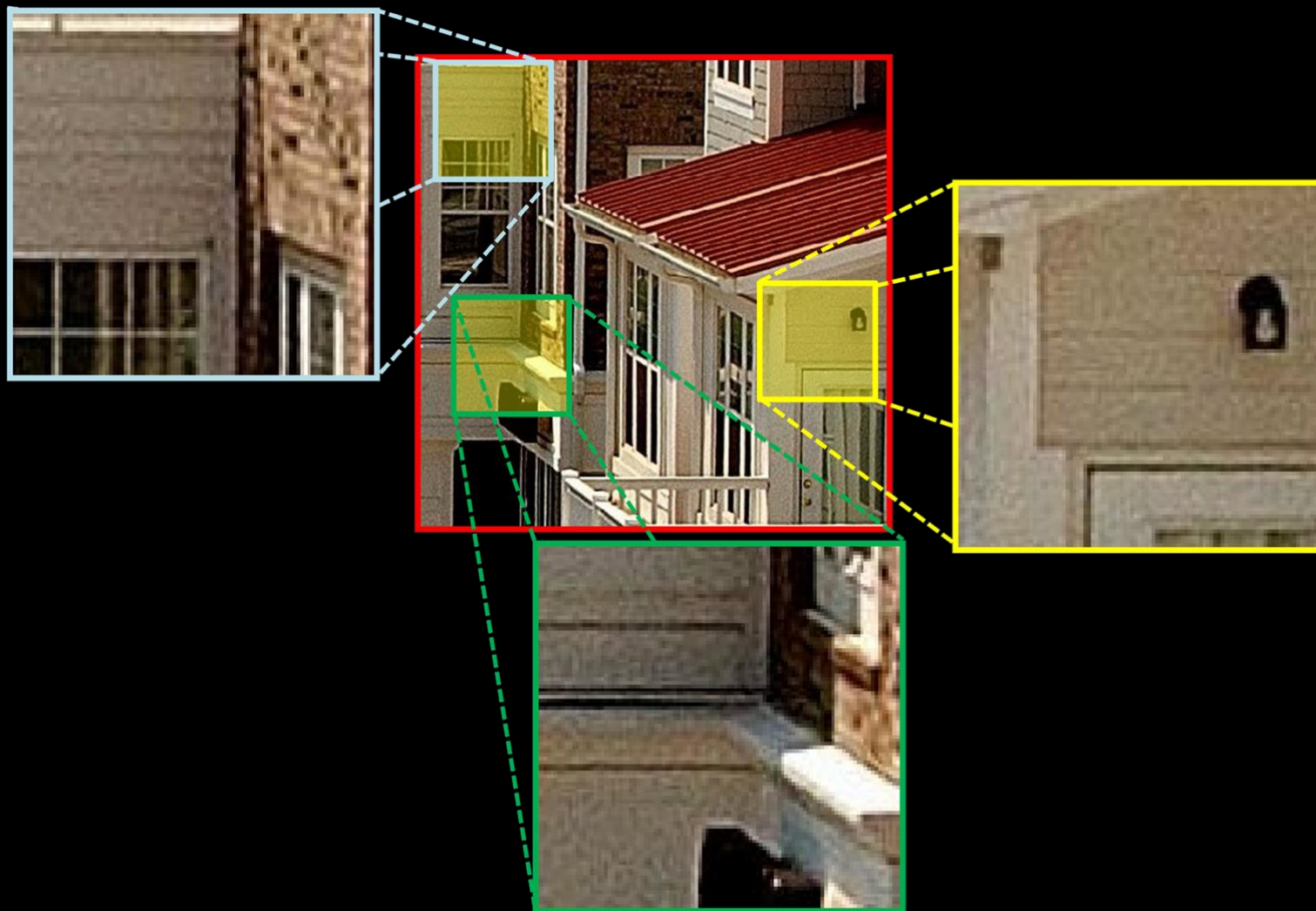
למה לעסוק בניקוי רעש? כי ...

- ❑ זו בעיה **אמיתית!** כל מצלמה דיגיטלית זקוקה לפתרון לסוגיה זו
- ❑ ניקוי רעש זו אחת הבעיות הבסיסיות ביותר בעיבוד תמונות ולכן מהווה **פלטפורמה מצויינת** על מנת לבחון רעיונות חדשים בתחום זה
- ❑ מסתבר שאם אנו יכולים לנקות רעש מתמונה, ניתן **למנף** זאת למשימות אחרות חשובות (חידוד, דחיסה, שחזור, ועוד ...)

# הצצה מקרוב – מחקר?



# הצצה מקרוב – מחקר?





# הצצה מקרוב – מחקר?

☐ הגישה הקלאסית בה פעלנו בשנים 1960-2010: בניית מודל מתמטי

# למבניות החבוייה בתמונה ומינוף

## מודל זה לבניית אלגוריתם לניקוי

□ למה הכוונה "מודל מתמטי" ?

# על איזו מבניות מדובר ?

□ נתחיל עם מהותה של תמונה:

- תמונה אינה אלא טבלה של מספרים שמייצגים "רמת אפור" בכל פיקסל

■ ואכן, רואים כי אזורים בהירים

הינם בעלי ערך גבוה הקרוב  
ל-100 ואזורים כהים קרובים ל-0.

■ **אגב, כשמחשב מקבל תמונה**

ל-"עיבוד", הוא למעשה פועל

## על מספרים אלו

|          |        |        |         |        |        |        |        |        |            |        |        |        |        |        |                                                               |
|----------|--------|--------|---------|--------|--------|--------|--------|--------|------------|--------|--------|--------|--------|--------|---------------------------------------------------------------|
| 11080809 | 122010 | 172739 | 344840  | 192015 | 1512   | 1821   | 212011 | 133230 | 171516     | 1621   | 1617   | 202123 | 203232 | 233221 | 212421                                                        |
| 10081011 | 100911 | 172053 | 717643  | 392317 | 1417   | 2123   | 232221 | 182755 | 171817     | 181717 | 181718 | 182242 | 245224 | 243438 | 382328                                                        |
| 10081119 | 1213   | 152519 | 3775650 | 241815 | 1414   | 2525   | 252424 | 274850 | 23202027   | 312025 | 532632 | 447232 | 285572 | 47649  |                                                               |
| 12111221 | 152023 | 232829 | 614357  | 372018 | 1820   | 2522   | 252624 | 246068 | 2120220222 | 262323 | 352229 | 697661 | 317276 | 707663 |                                                               |
| 34422023 | 212623 | 263930 | 577955  | 222830 | 242222 | 172627 | 223530 | 352222 | 222324     | 242423 | 233969 | 6163   | 557367 | 737673 |                                                               |
| 716827   | 292525 | 2731   | 5035    | 4271   | 3562   | 4924   | 2222   | 202727 | 2656       | 2424   | 5224   | 2524   | 253441 | 2625   | 26334550566359627573                                          |
| 634934   | 4725   | 2630   | 3051    | 4053   | 5866   | 7843   | 3125   | 2829   | 2426       | 2824   | 5529   | 4053   | 2528   | 2925   | 3169282625364040455450637573                                  |
| 574139   | 3831   | 3333   | 3945    | 5861   | 7179   | 8079   | 7332   | 3130   | 2929       | 2828   | 5856   | 7849   | 2628   | 2826   | 2839412728283835364746426364                                  |
| 474336   | 0433   | 3736   | 4243    | 4653   | 5563   | 7078   | 6551   | 5951   | 3132       | 3830   | 3067   | 7949   | 2733   | 3127   | 332730273229342926293329404                                   |
| 374239   | 3521   | 2635   | 3741    | 4047   | 4244   | 4550   | 5586   | 6674   | 6940       | 4928   | 3230   | 4067   | 8395   | 5925   | 45292931323337393626262529282636                              |
| 533938   | 2829   | 3138   | 3423    | 2733   | 3339   | 3948   | 4643   | 4951   | 5539       | 6232   | 3431   | 3146   | 6379   | 6564   | 522930322947635031302936282529                                |
| 375962   | 6156   | 4644   | 4032    | 3939   | 3930   | 2929   | 3237   | 4647   | 4142       | 5037   | 4542   | 3348   | 5067   | 6440   | 313024239666342453739322526                                   |
| 496053   | 4348   | 6564   | 6357    | 5044   | 3839   | 4330   | 3330   | 2634   | 4343       | 3524   | 3544   | 5566   | 6360   | 4539   | 282121706965584136372214                                      |
| 496151   | 4243   | 4664   | 4450    | 5056   | 6364   | 5751   | 2414   | 4338   | 3633       | 2633   | 3837   | 4442   | 4143   | 4857   | 3837303055697070666566463513310                               |
| 456525   | 2752   | 6444   | 4504    | 4249   | 6356   | 6265   | 0354   | 5451   | 4432       | 3028   | 3738   | 3842   | 6048   | 5665   | 70707171716566666664646360                                    |
| 616465   | 6565   | 6464   | 6054    | 4946   | 5059   | 4558   | 6149   | 5461   | 6665       | 5934   | 4041   | 3739   | 3444   | 6667   | 7171727171716758616364636360                                  |
| 596243   | 4555   | 6463   | 6666    | 6665   | 6454   | 5359   | 4643   | 4444   | 5046       | 5766   | 6053   | 5159   | 6971   | 7172   | 72717163514738404554606259                                    |
| 566146   | 5459   | 4342   | 6256    | 5865   | 6767   | 6665   | 6356   | 5244   | 4237       | 4254   | 8667   | 7070   | 7172   | 7272   | 7268564749585946444239384146                                  |
| 606359   | 5475   | 5745   | 4464    | 4946   | 5647   | 5459   | 6762   | 6566   | 752        | 3438   | 4768   | 7071   | 7272   | 7270   | 6351505595339292232934404240                                  |
| 636464   | 6565   | 6464   | 6166    | 5348   | 5140   | 4653   | 5735   | 3259   | 6971       | 7273   | 7369   | 5650   | 5361   | 5645   | 2934455446433326202329                                        |
| 615544   | 4864   | 6416   | 6364    | 6565   | 6460   | 4732   | 3739   | 5663   | 707170     | 707164 | 5351   | 5759   | 4833   | 3453   | 555555546474547474232                                         |
| 615546   | 4465   | 1437   | 6437    | 6355   | 5865   | 6655   | 4244   | 6436   | 707170     | 7066   | 5555   | 15361  | 5238   | 4254   | 57575655555343474647474474                                    |
| 635953   | 4864   | 5048   | 5262    | 4850   | 6146   | 4562   | 6769   | 707169 | 605050     | 5756   | 4330   | 3850   | 5557   | 5757   | 575253535447454243434545                                      |
| 596163   | 6566   | 6666   | 1616    | 1554   | 6555   | 5967   | 697070 | 6755   | 4953       | 53747  | 343444 | 5556   | 5655   | 5653   | 52535356565647484548474441                                    |
| 605266   | 5355   | 5849   | 5553    | 5962   | 6667   | 6869   | 7070   | 4058   | 5238       | 3334   | 4454   | 5352   | 5535   | 5355   |                                                               |
| 615755   | 4952   | 5048   | 5161    | 6166   | 6869   | 6969   | 697067 | 4734   | 1524       | 5455   | 5553   | 5053   | 5454   | 5556   | 7070706941514154525352545456555555                            |
| 554342   | 4650   | 5463   | 6667    | 6869   | 706969 | 707069 | 4151   | 4154   | 5253       | 5254   | 5456   | 5555   | 5555   | 5555   | 6970694239365153555455545656565555                            |
| 514033   | 4961   | 6667   | 6869    | 6868   | 6969   | 6969   | 6969   | 6969   | 4239       | 3651   | 53555  | 545554 | 5554   | 5556   | 5555                                                          |
| 515157   | 6566   | 6767   | 6867    | 6867   | 707069 | 707069 | 7072   | 6939   | 3434       | 5455   | 5455   | 5243   | 3030   | 3131   | 69706941514154525352545456555555                              |
| 475864   | 6566   | 6667   | 6666    | 6363   | 697071 | 707171 | 7172   | 6832   | 313046     | 3835   | 2820   | 1413   | 1114   | 1413   | 707071717172726728212319151216141413131511                    |
| 366165   | 6426   | 6362   | 6161    | 707071 | 717171 | 7272   | 6728   | 2123   | 1915       | 1216   | 1414   | 1313   | 1511   | 1413   | 707071717172726728212319151216141413131511                    |
| 316459   | 5961   | 6362   | 6360    | 707171 | 7172   | 7373   | 7266   | 2216   | 1412       | 1110   | 1111   | 1114   | 1719   | 1919   | 707071717172726728212319151216141413131511                    |
| 406458   | 6163   | 6464   | 6462    | 717271 | 717272 | 7265   | 2015   | 1513   | 2317       | 1617   | 181108 | 0811   |        |        | 7172726728212319151216141413131511                            |
| 416559   | 6363   | 6564   | 6462    | 707172 | 7272   | 6420   | 1717   | 1515   | 1078       | 1218   | 2530   | 4475   |        |        | 7172726728212319151216141413131511                            |
| 466261   | 6363   | 6564   | 6767    | 717373 | 737373 | 7262   | 1719   | 2331   | 4046       | 5055   | 2525   | 5048   | 464    |        | 7172726728212319151216141413131511                            |
| 496263   | 6567   | 717273 | 727272  | 7167   | 6359   | 3495   | 050504 | 5451   | 3333       | 3736   | 2841   |        |        |        | 7172726728212319151216141413131511                            |
| 516568   | 6971   | 7170   | 6662    | 5855   | 1474   | 4411   | 3939   | 432    | 4243       | 4740   | 4431   | 3244   | 49     | 49     | 516568697171706662585514744411393943242434740443132444949485  |
| 526058   | 5551   | 4645   | 4948    | 4953   | 1405   | 3532   | 3141   | 4951   | 6255       | 4344   | 3437   | 35148  | 4441   |        | 52605855514645494849535140353231414951625543443437514844414   |
| 331242   | 1183   | 3425   | 5153    | 5554   | 4034   | 3434   | 3947   | 5153   | 3394       | 4033   | 3848   | 4730   | 284    |        | 3312421183342551535558544034343439475153339403338484730284    |
| 322009   | 0810   | 3534   | 4336    | 3543   | 2834   | 3534   | 3533   | 1939   | 5052       | 5336   | 3631   | 3647   | 47     | 28     | 3220090810353443363543283435343533193950525336363136474728284 |
| 311411   | 1012   | 3224   | 3225    | 3041   | 2733   | 3331   | 3132   | 1738   | 4750       | 3636   | 3623   | 5444   | 43     | 35     | 3114111012322432253041273333313217384749503636235444335314    |
| 341517   | 1519   | 3534   | 4141    | 4245   | 4033   | 3029   | 2817   | 3644   | 4547       | 3747   | 3334   | 2421   | 12     | 26     | 34151715193534414142450433302928173644454737473334242112263   |
| 371413   | 1330   | 3834   | 3939    | 4148   | 4832   | 2625   | 2414   | 3338   | 3943       | 3933   | 2529   | 3136   | 16     | 24     | 37141313303834393941484832262524143338394339332529313616243   |
| 595152   | 5257   | 5854   | 5554    | 5266   | 4540   | 3835   | 3234   | 3435   | 3544       | 3426   | 3131   | 2321   | 31     |        | 59515252575854555452664540383532343435354434263131232131      |
| 798080   | 8080   | 8080   | 8081    | 8181   | 8181   | 8282   | 8282   | 8180   | 7978       | 7675   | 7372   | 706    |        |        | 798080808080808181818181828282828180797876757372706           |

|    |    |    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|----|----|
| 70 | 70 | 69 | 41 | 51 | 41 | 54 | 52 | 53 |
| 69 | 70 | 69 | 42 | 39 | 36 | 51 | 53 | 55 |
| 70 | 72 | 69 | 39 | 40 | 34 | 54 | 54 | 55 |
| 71 | 72 | 68 | 32 | 31 | 33 | 46 | 38 | 35 |
| 72 | 72 | 67 | 28 | 21 | 23 | 19 | 15 | 12 |
| 72 | 72 | 66 | 22 | 16 | 14 | 14 | 21 | 10 |
| 72 | 72 | 65 | 20 | 15 | 15 | 13 | 23 | 17 |
| 72 | 72 | 64 | 20 | 17 | 17 | 15 | 17 | 08 |



# הצצה מקרוב – מחקר?

☐ הגישה הקלאסית בה פעלנו בשנים 1960-2010: בניית מודל מתמטי

**למבניות החבוייה בתמונה ומינוף מודל זה לבניית אלגוריתם לניקוי**

□ למה הכוונה "מודל מתמטי" ?

# על איזו מבניות מדובר ?

☐ שימו לב: פיקסלים סמוכים נוטים להיות קרובים בערכיהם – זוהי

## מבניות שנקראות

# Piecewise-Smoothness

□ ניתן לנסח תכונה זו כביטוי

## מתמטי ו-"למדוד" כמה

## התמונה מקיימת זאת

11080809 122010 172739 344840 192015 15 12 1821 212011 1133230 171516 1621 1617 202123 202322 233221 2121 2421  
1008 10 11 1009 11 17 205371 7264 392317 14 1721 23232221 182755 171817 1817 1817 18 2224 24 5224 24 34 3838 2328  
1008 11 19 12 13 1525 1937 75765024 1815 14 14 2525 2524 23 27 4860 23202027 312025 532632 44 7232 2855 7247 6449  
12 11 1221 152323 232829 614357 37 2018 18 202522 252624 246068 21202222 262323 35222969 7661 31 72 7670 7663  
344220231 2623 263930 5271 7555 22 283024 2222 17 2627 22 35 3035 222223 22 2324 24 262339 6961 63 55 7367 73 7673  
7168272925252 315035 4271 35 42569422 222220 272726 5624 24 5224 2524 253441 242526 334550 56635962 7573  
63493447 2526 30305140 535866 784331 25282924 262824 552940 53252829 253169 282625 364040 4554 5063 7573  
57413938 313333 3945586174 79807973 32 31 30292828 5856 78 49262828 26283941 27282838 35 3647 4642 6364  
47434360 4337 36 424346 535563 707865 515951 31 322830 306779 49274331 27 23327 3029 3429 26293329 4048  
37423935 21 2635 374140 47 424550 525866 746940 4928 32 304678 535952 45 292931 323337 3936 26 2625 2928 2636  
35 393828 2931 38 342327 333939 48464349 51553962 3234 31 3463 63796564 522930 322947 635031 30293628 252929  
37596261 564644 403239 39302929 3237 464741 425037 45 423348 706767 644031 30 2420 39666342 45 373932 2526  
49605314 686564 635750 443839 433033 20364343 43352 543544 55626360 453928 21 2134 706965 58413637 22 14  
49615142 436644 4450566364 57514241 43 3836 33263338 3744241 434857 3837 30 305569 707066 65646351 3310  
45626257 526444 454042 42496356626560 534541 1444 32 302837 38384260 485665 707071 717165 66666664 6458  
61646565 656464 605449 465059454861 49546166 595343 302137 3934 4466677171 727171 716758 616364636360  
59624345 5564366666666656454 535946 43444 4504657 666053 51596971 717272 7217635147 3840 455460 6259  
56614646 545943 426556 58656767666563 565244 423742 654866 707071 72727268 564749585946 444239384146  
60635954 57574546649 466575495762656657 243847 687071 727272 706351 50 555953 395922 2329340 4240  
636464 65656464 616653 486053 48510406 5357 35 3259696971 727273695650 536156 452934 454423 332620 232969  
61554448 646163646465 65656460 473237 395663 707170 707164 535157 59483333 455355 555464 4754747474232324745  
61554644 65514347363558646566554244 647070 71706655515364 55123829 425457 575655 555343474746474745  
63595348 645048526248 50614645626769 707169605050 575643 303850555757 5757 575253 535447 45424343 4545  
596163656666616155 46555967697070675549535747 343444 55565655565352 535356565647 485448474441

|    |    |    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|----|----|
| 70 | 70 | 69 | 41 | 51 | 41 | 54 | 52 | 53 |
| 69 | 70 | 69 | 42 | 39 | 36 | 51 | 53 | 55 |
| 70 | 72 | 69 | 39 | 40 | 34 | 54 | 54 | 55 |
| 71 | 72 | 68 | 32 | 31 | 33 | 46 | 38 | 35 |
| 72 | 72 | 67 | 28 | 21 | 23 | 19 | 15 | 12 |
| 72 | 72 | 66 | 22 | 16 | 14 | 14 | 21 | 10 |
| 72 | 72 | 65 | 20 | 15 | 15 | 13 | 23 | 17 |
| 72 | 72 | 64 | 20 | 17 | 17 | 15 | 17 | 08 |

# הצצה מקרוב – מחקר?

□ הגישה הקלאסית בה פעלנו בשנים 1960-2010: בניית מודל מתמטי

למבניות החבויה בתמונה ומינוף  
מודל זה לבניית אלגוריתם לניקוי

□ למה הכוונה "מודל מתמטי" ?

על איזו מבניות מדובר ?

□ שימו לב: אירועים בתמונה נוטים

לחזור על עצמם – זוהי מבניות

אחרת הנקראית **Non-Local**

**Self-Similarity**

□ ניתן לנסח גם תכונה זו כביטוי

מתמטי ו-"למדוד" כמה התמונה

מקיימת אותה





# הצצה מקרוב – מחקר?

□ ברוח זו ניתן לזהות תכונות מובניות נוספות בתמונות – נזכיר ללא פירוט:

ניתן לנסח כל תכונות אלו  
כביטויים מתמטיים ו-"למדוד"  
כמה התמונה מקיימת אותן

- Scale Invariance
- Sparse Representation
- Low-dimensionality

□ ואז, בהינתן תמונה רועשת, איך נוכל לנקותה?

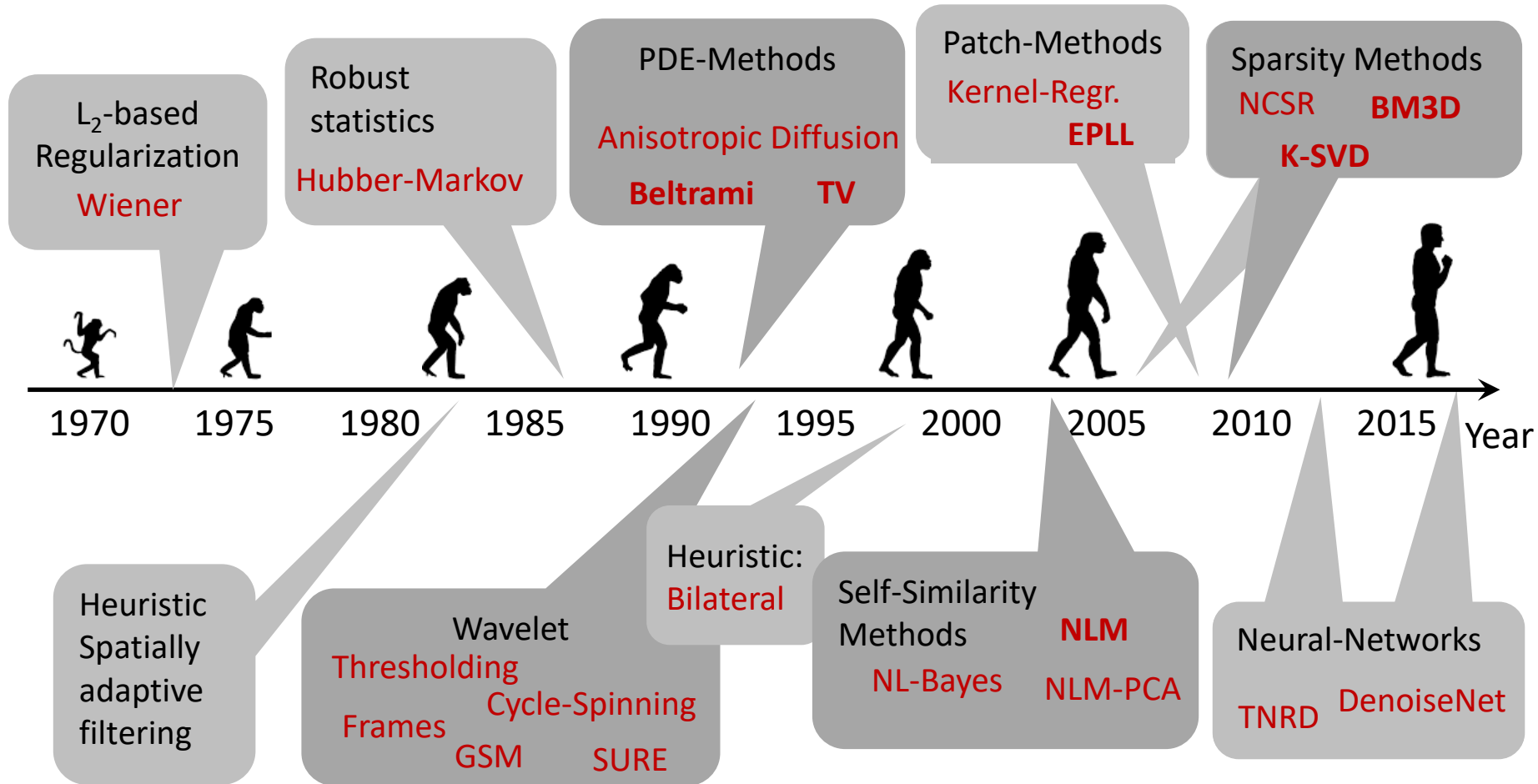
□ ... ננסח תהליך מתמטי ש-"ימשוך אותה" לקיום טוב יותר של כל התכונות הללו, וכך ניפטר מהרעש

□ יתרה מזו: אנו אפילו מסוגלים לנתח תיאורטית ביצועים של גישות אלה

□ ברוח זו פותחו שפע אלגוריתמים יפיפיים שתרמו להתפתחות התחום הקרוי "עיבוד תמונה", ואיתם פותחה גם תיאוריה עמוקה ומרשימה

□ ללא הגזמה: יש כ- 20,000 מאמרים שעסקו בתחום והציעו מגוון פיתוחים ופתרונות שונים ומגוונים לניקוי רעש

# הצצה מקרוב – מחקר?



# הצצה מקרוב – מחקר?

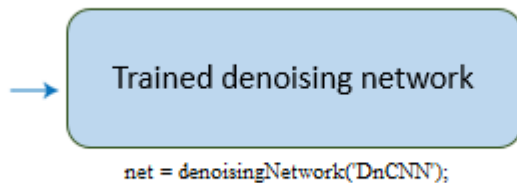
ואז הגיעה למידה עמוקה  
וטרפה את הקלפים

הגישה העכשווית פועלת כך:

- אסוף המון (נניח 10,000) צמדים של תמונות רועשות וגירסתן הנקייה
- הגדר ארכיטקטורה של רשת (רצוי עמוקה) שכניסתה ומוצאה הן תמונות
- אמן את הרשת לקבל תמונה רועשת ולהפיק ממנה תמונה שתהיה קרובה ככל האפשר לגירסתה הנקייה



Noisy Image



Denoised Image

□ ו... זהו! **זזה עובד**  
**יותר טוב מכל שיטה**  
**שפיתחנו בעבר**

# הצצה מקרוב – מחקר?

ומה עם כל הידע שצברנו לאורך השנים?

מה עם השיטות הקלאסיות שפיתחנו?

המודלים שהרכבנו?

התיאוריה המפוארת שיצרנו?

שימו לב שבגישה החדשה אין לכאורה מקום לכל אלו!

האם זה ברור מאליו?

האם כל הידע שיצרנו ירד לטמיון?

והנה שאלה יותר מרכזית: מה תכליתן האמיתית של עבודות מחקר: פתרון בעיות? השגת הבנה עמוקה יותר? משהו אחר?

Noisy Image

Denoised Image

# הצצה מקרוב – מחקר?

אז מה אני (ואחרים) עושים בכל זאת כדי למזג את הידע הרב מהעבר וגישת הלמידה החדשה שכתובת כל פיסגה?

מציאת מעקף לצורך  
בריבוי נתונים לאימון

פיתוח טכניקות  
אימון חדשות  
ושונות מהמקובל

בחירה ויצירה של  
ארכיטקטורות חדשות  
מושכלות

הבאת הבנה תיאורית  
של ביצועי רשתות בעזרת  
חקר קלאסי

הזרקת הבנה מודלית  
לתהליך האימון  
והפעלה לשיפור  
ביצועים

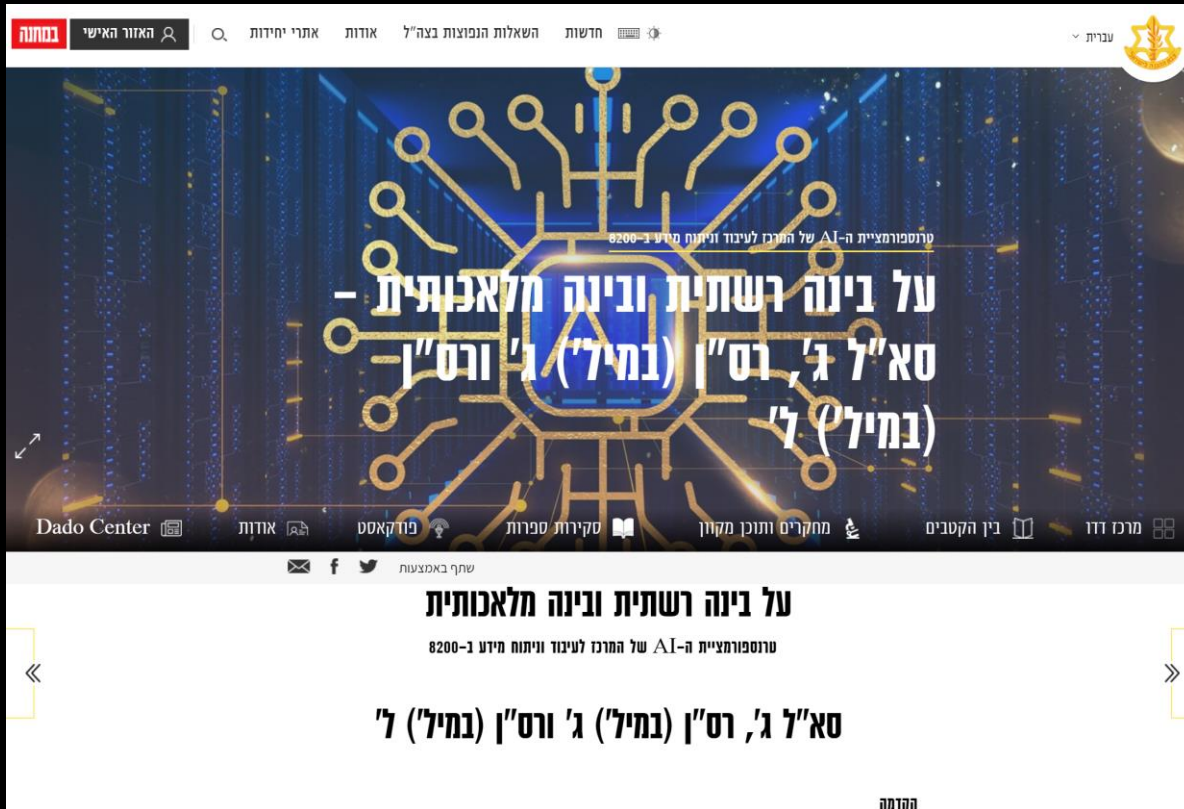
פתרון בעיות חדשות בעזרת  
למידה עמוקה עם תובנות  
מהקלאסיקה



# ומה עם הצבא?

צבאות מתקדמים בעולם הבחינו בזירה מתפתחת זו והתאימו עצמם אליה בשפע של דרכים:

- ☐ ראשי ביות לטילים מבוססי למידה עמוקה
- ☐ עקיבה ויירוט ע"י רחפנים וכתב"מים
- ☐ מערכות תצפית אוטומטיות
- ☐ ניהול מידע מודיעיני והפקת תובנות ממנו
- ☐ ועוד ועוד ...



כתבה מתחילת מרץ 2021 במרכז דדו

מעט על

אתגרים עתידיים

(חברתיים וטכנולוגיים\מדעיים)

# לאן מכאן ?

המדענים: עובדים על הבנת המכונה הזו יותר טוב, פיתוח  
**תיאוריה צלולה ומלאה** שתבהיר למה זה עובד  
ובאיזה תנאים, וחיפוש הרחבות לתחום זה

התעשייה: צועדת בדרך האחת והיחידה שעומדת לרשותה –  
**לעשות כסף** הכי קל והכי מהר עם ניצול הטכנולוגיה  
בצורה מיטבית למימוש מוצריה

רגולציה: כמו כל טכנולוגיה, ניתן למנף את היכולות החדשות  
שלמידה עמוקה מביאה על מנת להרע – **מי יעצור**  
את המדענים והמהנדסים מלהמציא כלים להרע  
לאנשים?

# לאן מכאן ?

וביתר פירוט, בואו נדבר בקצרה על ...

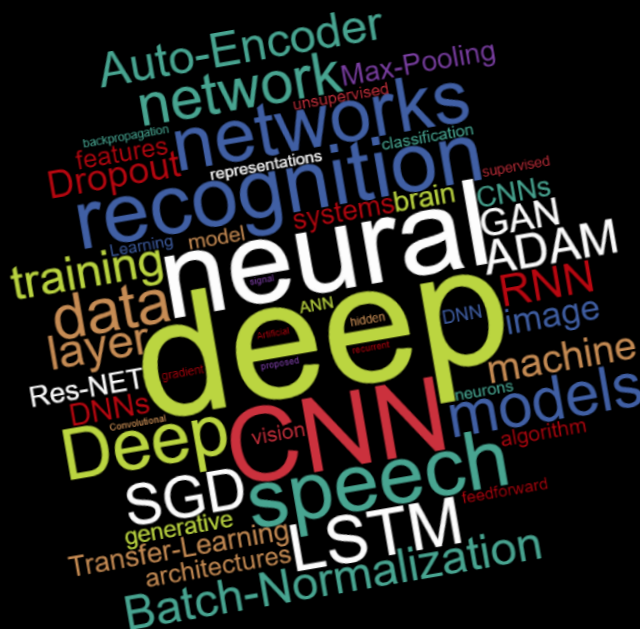
- ☐ צורך בהסבר תיאורטי
- ☐ רגישות להתקפות
- ☐ Deep Fake
- ☐ חשש להטייה
- ☐ בעיות מוסריות
- ☐ הנמקת החלטות
- ☐ לימוד בתנאי חוסר מידע
- ☐ בחירת הארכיטקטורה

# צורך בהסבר תיאורי

- אין ספק שהתחום "למידה עמוקה" הינו סיפור הצלחה יוצא דופן, עם יכולת חסרת תקדים בפתרון שורה ארוכה של בעיות הנדסיות
- אבל, תחום זה הוא אמפירי ביסודו, עתיר טריקים שאינם מובנים עד סופם (מגיה שחורה), ומבוסס בעיקר על ניסוי ותהיה\טעיה
- מה לא מובן לנו?

- איך יש לבחור ארכיטקטורה?
- מה גבולות היכולת של רשתות אלו?
- מה לא ניתן ללמוד?
- למה הרשת מצליחה ללמוד למרות ריבוי הפרמטרים?
- מה יכולנו לשפר ולעשות אחרת?

- מדענים רבים מנסים להביא הבנה תיאורטית לתחום זה במגוון אופנים – זהו ה- Holly Grail של התחום

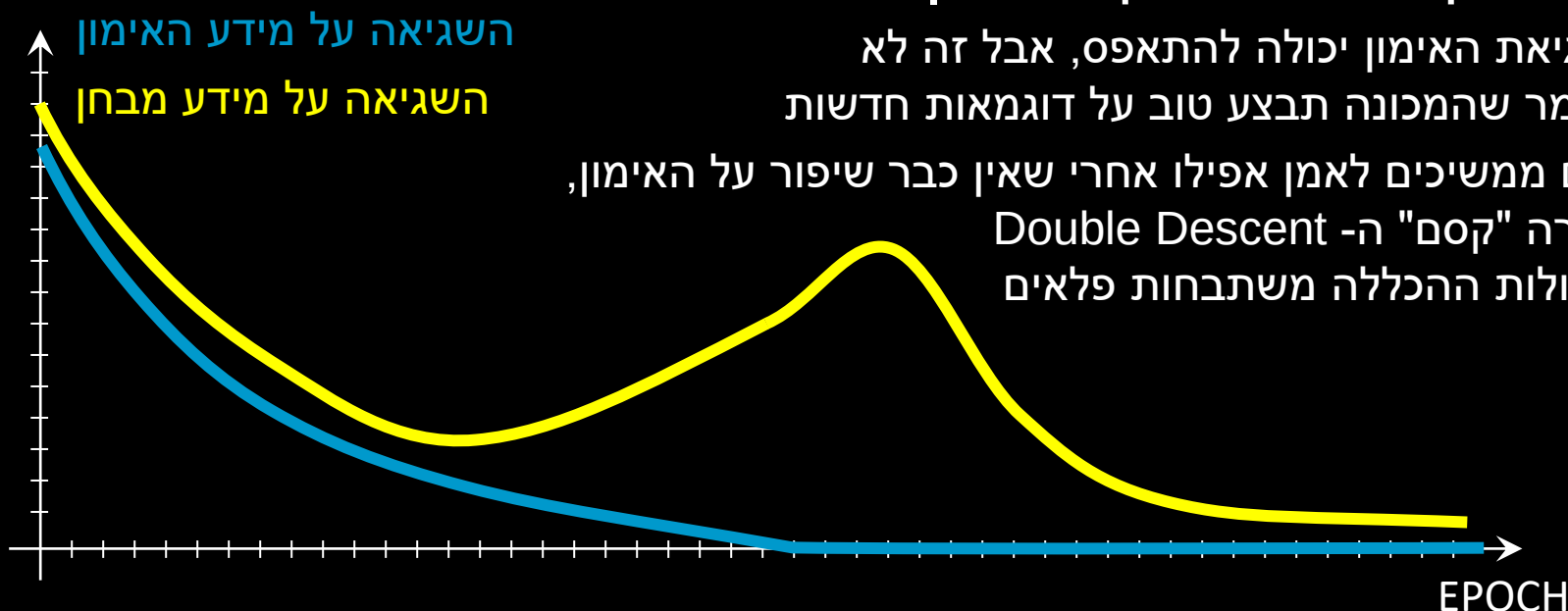


# צורך בהסבר תיאורטי

## Over-Parametrization and Double-Descent

□ נניח שעבור משימתנו אנו מאמנים רשת עם 10 מיליון פרמטרים, ויש לנו 10,000 דוגמאות לאימון

□ הנה – כך ייראה תהליך האימון:





# צורך בהסבר תיאורי

הנה סיפור שימחיש היטב את חסרונה של התיאוריה  
והתחושות הקשות בקרב חוקרים בקהילה כתוצאה מכך

Ali Rahimi:  
NIPS 2017  
Test-of-Time  
Award



“Machine  
learning  
has  
become  
alchemy”



Yan LeCun



Understanding is a good thing ... but another goal is inventing methods. In the history of science and technology, engineering

preceded theoretical understanding:

- Lens & telescope → Optics
- Steam engine → Thermodynamics
- Airplane → Aerodynamics
- Radio & Comm. → Info. Theory
- Computer → Computer Science

# צורך בהסבר תיאורטי

כאמור, שפע של חוקרים (מתמטיקאיים למיניהם) עובדים על בניית הנמקה תיאורטית ללמידה עמוקה

תופעה מעניינת בזירה זו – ריבוי השפות והכלים בהם נעשה שימוש

Optimization (Vidal, JHS), Signal Processing (Baraniuk, Rice),  
Control Theory (Soato, UCLA), Information Theory (Tishby, HUI),  
Harmonic Analysis (Mallat, ENS), Sparse Representation (me),  
PDE (Osher, UCLA), Machine learning (Poggio, MIT),  
Theoretical CS (Arora, Princeton), Neuroscience (Simoncelli, NYU), ...

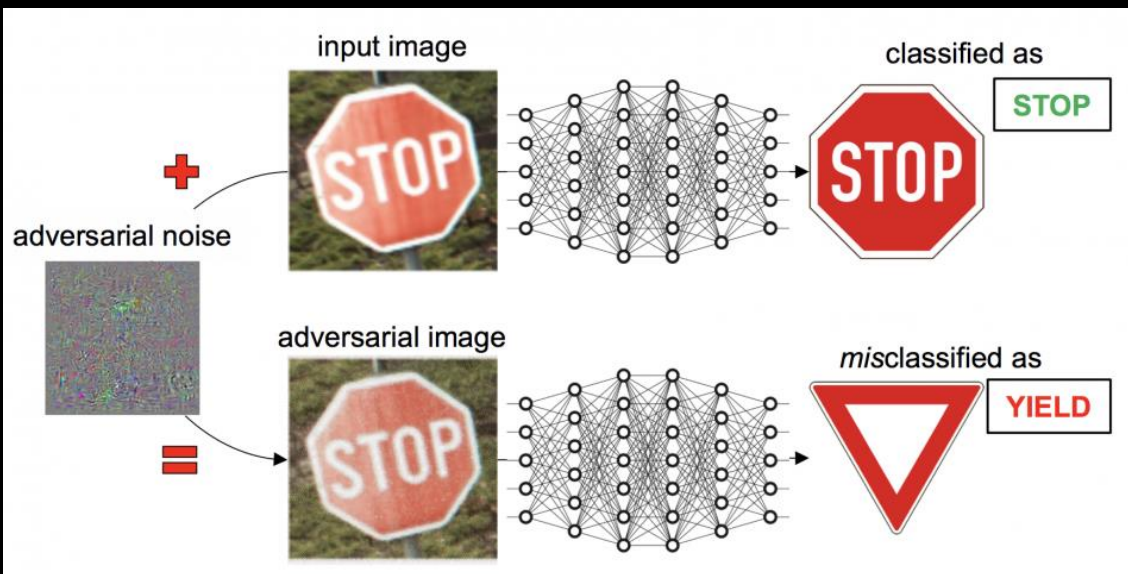


Ron Kimmel: *"DL is a dark monster covered with mirrors. Everyone **sees his reflection** in it ..."*



David Donoho: *"... these mirrors are taken from Cinderella's story, telling each that he is the **most beautiful**"*

# רגישות להתקפות

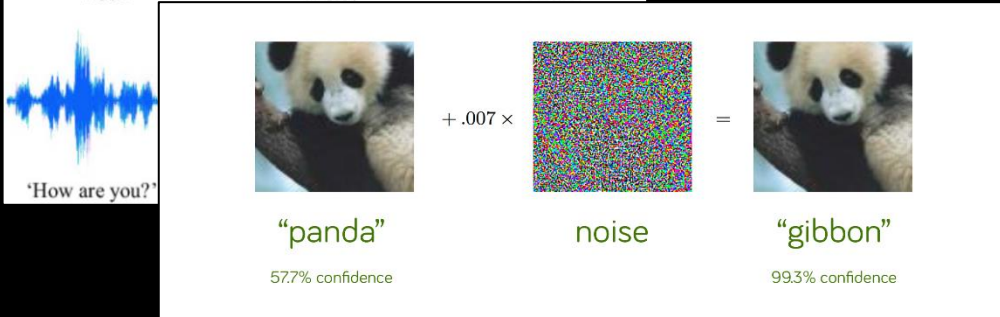
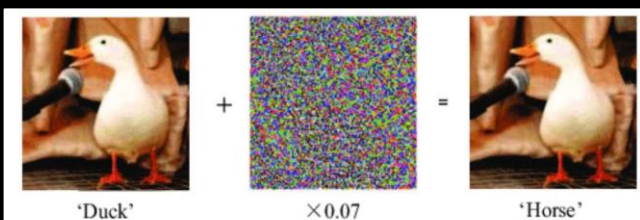


□ ב-2014 גילו חוקרים שרשתות עמוקות שאומנו (בהצלחה!) למשימותיהן מגלות רגישות בעייתית להתקפות רעש בכניסתן

□ סטיה קלה ומכוונת במידע בכניסה לרשת יכולה

לשבש לחלוטין את החלטותיה

□ תגלית זו הובילה לפעילות ענפה של בניית שיטות הגנה והתקפה על רשתות, והותירה אותנו עם חשש שאי-אפשר לסמוך על רשת מאומנת גם בשל סיבה זו



# Deep Fake

- ❑ מדובר באחת הרעות החולות של תחום הלמידה העמוקה – היכולת לזייף תכני אודיו, תמונות ווידאו
- ❑ לשימושים של יכולת זו אין גבולות ברורים, ולצד ההומור הנלווה, יש גם נזקים מעוררי חלחלה כגון הפצת מידע מטעה (fake news)



העתקת הבעות פנים מפרצוף לפרצוף



הכנסת פרצופו של אדם א'  
(Wolfram) לציור של אדם ב' (ביוטון)

# Deep Fake

האפקטים המרשימים ביותר הם סירטוני וידאו שהם הדובר  
לכאורה מצולם אך אומר דברים שמעולם לא הגה מפיו



הכנסת פרצופו של  
ניקולס קייג' לסרטו  
בו מדברת שחקנית  
אחרת

ויש עוד שפע של סרטונים כאלו  
(המפורסמים שבהם מציגים את  
אובמה, טראמפ, שוורצנגר,  
ואחרים) – חפשו ב- YouTube  
תחת Deep Fake



# Deep Fake

האפקטים המרשימים ביותר הם סירטוני וידאו שהם הדובר  
לכאורה מצולם אך אומר דברים שמעולם לא הגה מפיו

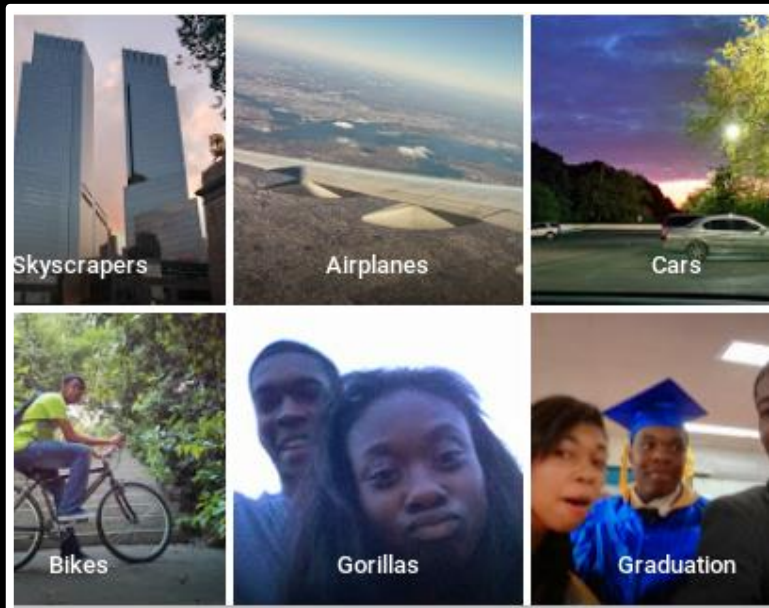


(מבית סאות' פארק) Sassy Justice הסדרה הסאטירית  
נעשית ע"י טכנולוגיית דיפ פייק. כוכב הסדרה הוא הבלש סאסי  
(דונלד טראמפ), והוא מראיין דמויות סינתטיות של מפורסמים

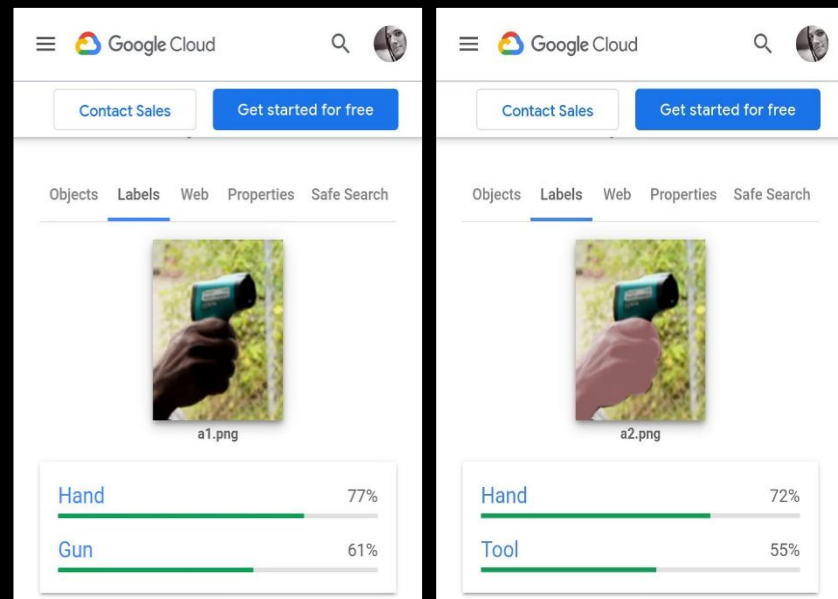


# חשש להטייה

- ❑ עם כניסתן של רשתות עמוקות למגוון תחומים, התגלתה נטיה בעייתית נוספת שלהן – חשש להטייה (bias) בהחלטות הרשת
- ❑ הנטיה הזו לא צריכה להפתיע – כיוון שהלמידה נשענת על מאגרי מידע מאסיביים, אם אין איזון בין הדוגמאות השונות, הרשת תתמיע זאת
- ❑ דוגמאות מפורסמות שחשפו בעיה זו גרמו למבוכה גדולה לגוגל



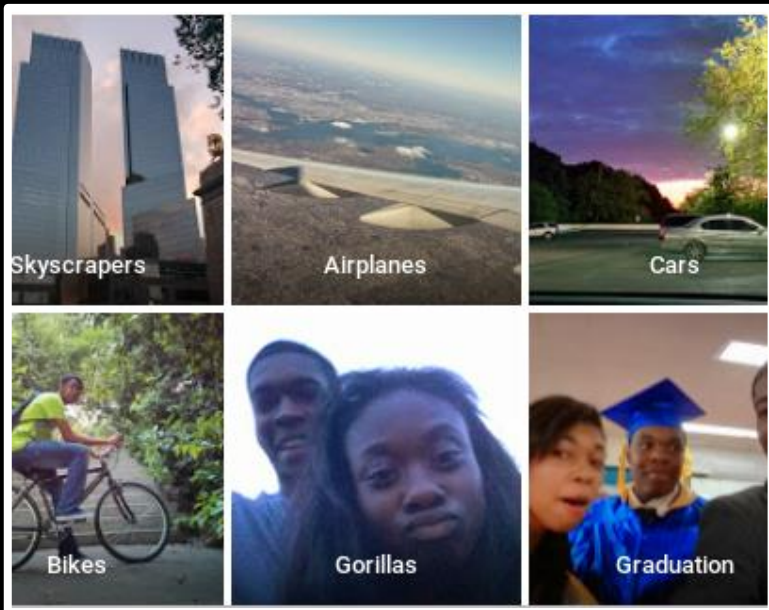
שירות זיהוי אוטומטי ב- Google Photos הצית אש ביוני 2015 בשל תוצאות הזיהוי בתמונות אלו



באפריל 2020 שוב חזרה התופעה והפעם בכלי זיהוי תמונה של שירות Google Cloud

# חשש להטייה

- ❑ עם כניסתן של רשתות עמוקות למגוון תחומים, התגלתה נטיה בעייתית נוספת שלהן – חשש להטייה (bias) בהחלטות הרשת
- ❑ הנטיה הזו לא צריכה להפתיע – כיוון שהלמידה נשענת על מאגרי מידע מאסיביים, אם אין איזון בין הדוגמאות השונות, הרשת תתמיע זאת
- ❑ דוגמאות מפורסמות שחשפו בעיה זו גרמו למבוכה גדולה לגוגל



- ❑ ב-2015 התברר שרשת עמוקה לסינון מועמדים באמזון "**מעדיפה גברים**". נסיון לסלק מהמאגרים את איזכור המין של המועמד לא צלח – הרשת ניסתה לנבא את מין המועמד והמשיכה בשלה
- ❑ גם כאן, פעילות ענפה של מחקר ופיתוח נוצרו על מנת להתמודד עם אתגר זה

שירות זיהוי אוטומטי ב- Google Photos הצית אש ביוני 2015 בשל תוצאות הזיהוי בתמונות אלו

# ועוד בעיות מוסריות



□ ניתן לאסוף המון (עשרות אלפי ויותר) דוגמאות של תמונות פנים

- של עבריינים ואנשים "רגילים"
- של הומו- והטרו-סקסואלים
- וכנגדם ציון פסיכומטרי או IQ
- ונגדם מצב בריאותי

□ בכל מקרים אלו, נוכל לאמן רשת לגלות את הקשר בין הפרצוף ובין מאפיינים אלו, ואז להפעיל זאת ככלי החלטה מול אנשים חדשים

□ באופן כללי יותר, לטכנולוגיית הלמידה העמוקה גם פוטנציאל להרע: כלים אלו כבר משמשים לעיבוד מידע שמוקלט בהחבא מהטלפון שלכם ולרגל אחריכם. טכנולוגיה זו משמשת למעקב אחר אזרחים במדינות מסויימות

# הנמקת החלטות

□ אנו יכולים כיום לאמן רשתות למשימות מגוונות ולצפות לדיוקי החלטה גבוהים ... אבל החלטות אלו באות ללא הנמקה

□ כשמדובר בבידור או בנושא קליל אחר, זה לא חשוב ... אבל מה אם

■ הרשת איבחנה מחלה והגדירה תהליך טיפולי?

■ הרשת החליטה שאינך זכאי להלוואה בבנק?

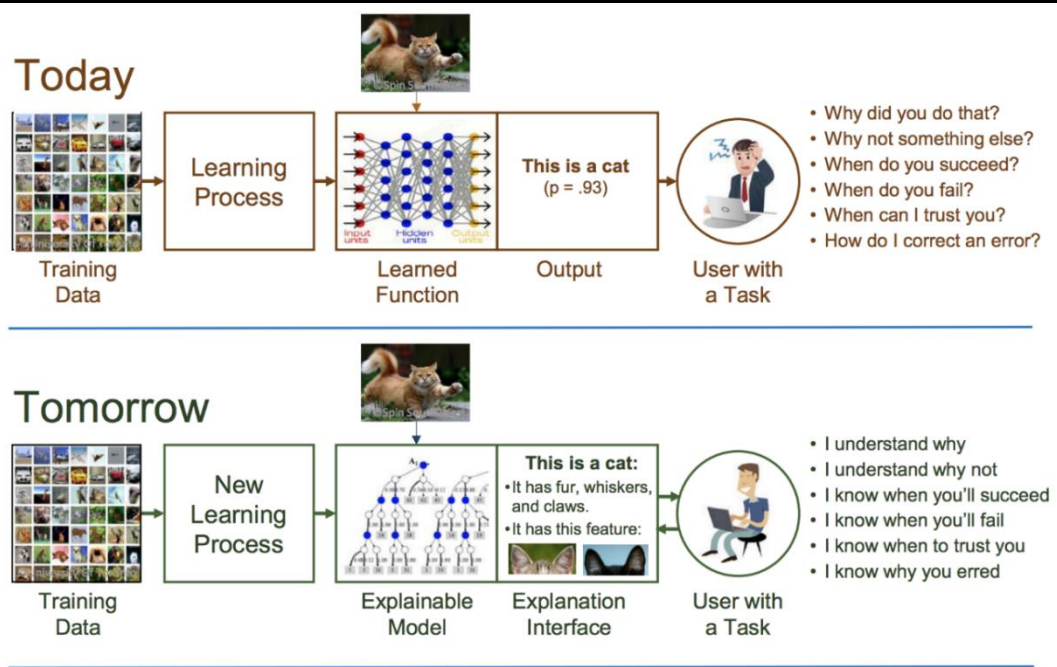
■ הרשת החליטה לסובב

את ההגה שמאלה

בחדות והביאה לתאונה?

■ הרשת מנעה ממך ויזת שהיה

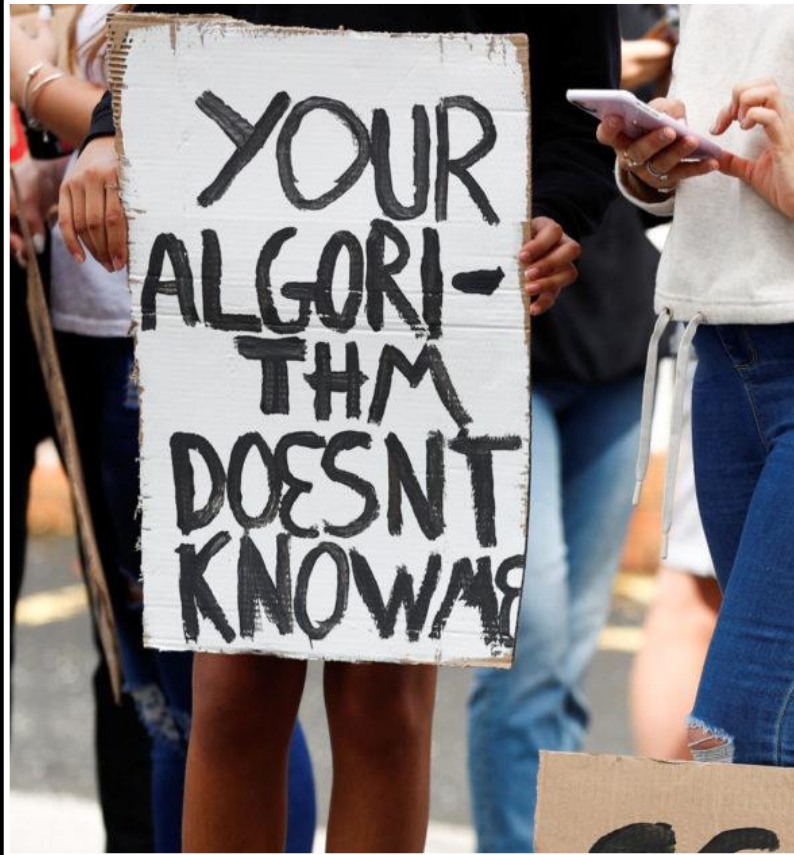
במדינה בה אתה מתארח?





# הנמקת החלטות

- אנו יכולים כיום לאמן רשתות למשימות מגוונות ולצפות לדיוקי החלטה גבוהים ... אבל החלטות אלו באות ללא הנמקה
- כשמדובר בבידור או בנושא קליל אחר, זה לא חשוב ... אבל מה אם



- הרשת קבעה את דירוג האשראי שלך?
- הרשת קבעה אם תתקבל לאוניברסיטה?
- ... העבודה על סוגיה זו בעיצומה ...

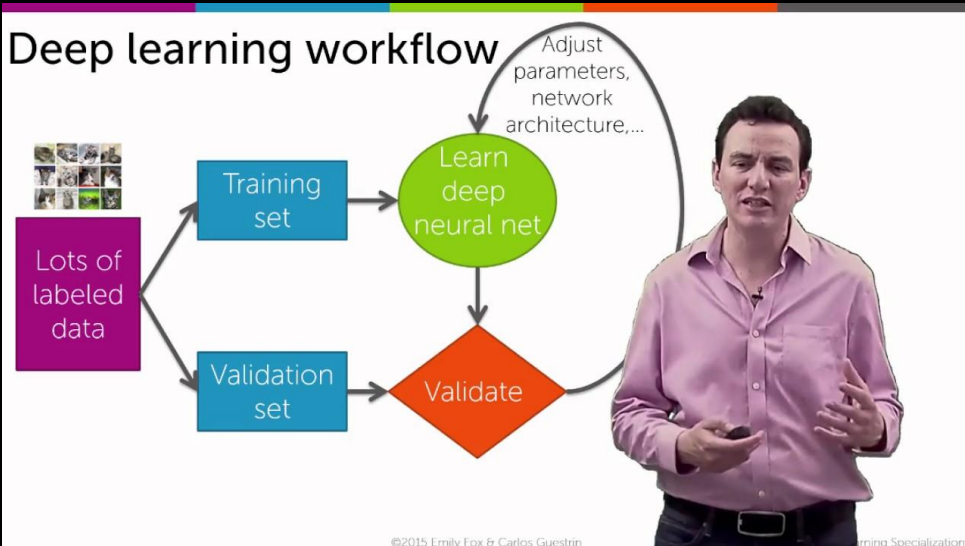
# לימוד בתנאי חוסר מידע

□ הגישה הרווחת בלמידה עמוקה נשענת על איסוף מאגרי מידע גדולים ומתויגים ואימון רשת עליהם – גודל המאגר עליו מתאמנים קריטי להצלחה

מה עושים כשאינ מספיק מידע? האם למידה עמוקה עדיין אפשרית?

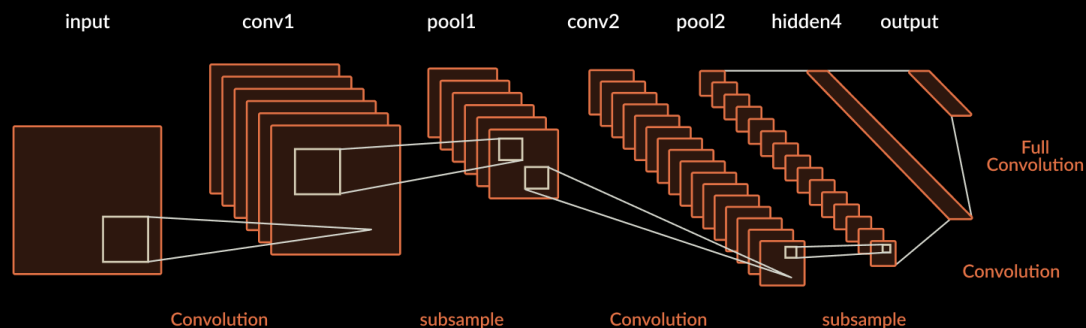
למשל: יש לנו שפע עצום של צילומי CT ו-MRI, אבל אין לנו תיוג של האיזורים הבעייתיים ואיפיונם ע"י רדיולוגים

שפע של עבודות מחקר תוקף  
סוגיה זו ומציע מגוון טכניקות  
של "לימוד לא מונחה" או  
"לימוד עצמי"





# בחירת הארכיטקטורה



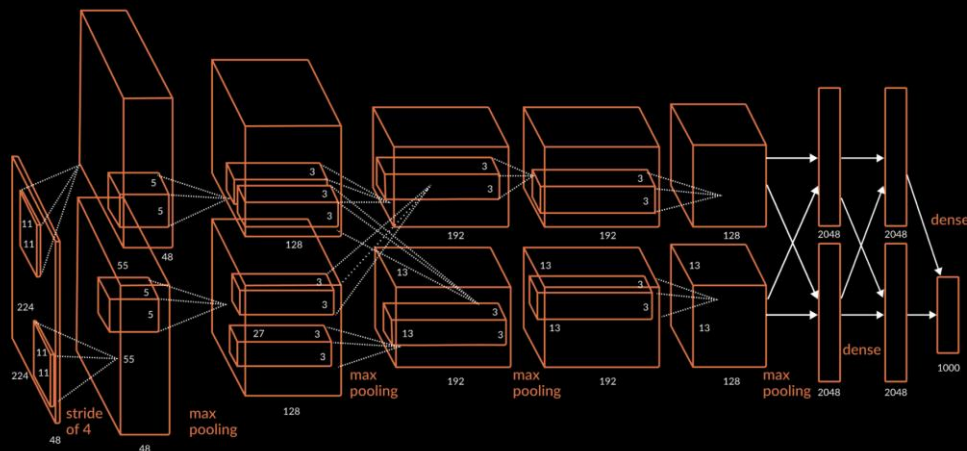
□ הנה משימה: בנה מכונה שתגלה אוטומטית טנקים בתמונות אויר ותסמן. נניח כי לצורך זה נאסוף אלפי תצ"אות וסימוני מטרות עליהן

□ נהדר, בוא נאמן את רשת עמוקה ונייצר תהליך אוטומטי

ואמין של ניתוח הדמאות, שיוכל להחליף חיילים שעוסקים בפענוח

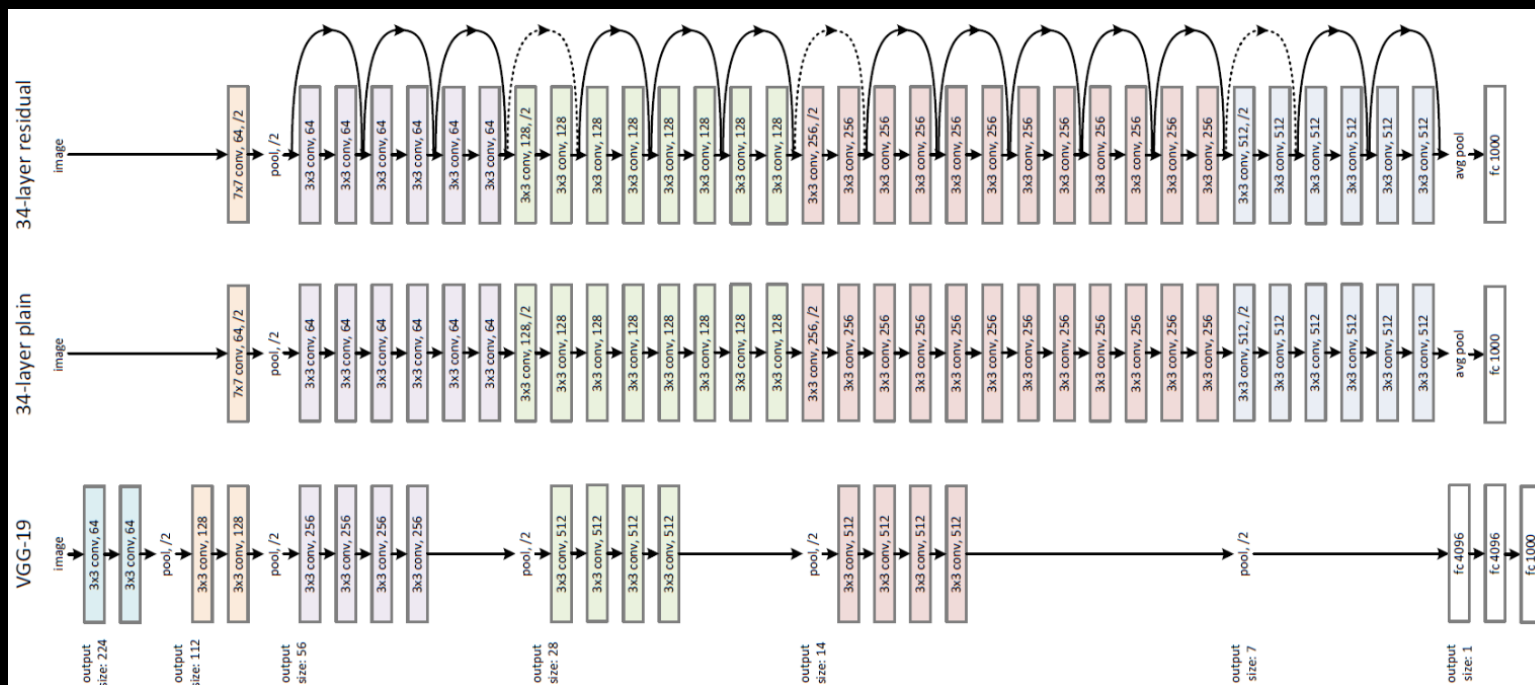
□ רגע ! באיזו רשת נשתמש?

מה תהיה צורתה? כמה נירונים  
יהיו בה? כמה שכבות? האם  
למנף כלים מוכרים כמו BN,  
Skips, Leaky-ReLU  
? ... Pooling, Stride,



# בחירת הארכיטקטורה

- אז לידיעתכם: בניית ארכיטקטורה של רשת כיום נעשית ע"י ניחוש ולא העתקה.
- מה עם דרך מושכלת לבחירת הארכיטקטורה?
- תשובה אפשרית טמונה בנושא שהעלינו קודם – מינוף הקלאסיקה על מנת להציע רשתות מונמקות
- תשובה אחרת: NAS – Neural Architecture Search



# אתגרים: לסיכום

לצד ההצלחה הכבירה של למידה עמוקה בשורה ארוכה של אתגרים הנדסיים, יש משימות לא פשוטות לפנינו:

- **פרצות הנדסיות שמחייבות מענה** – רגישות להתקפות, נוחות בייצור זיופים, עבודה עם מידע חסר, בחירת ארכיטקטורה, חשש להטייה
- **אתגרים מחקרניים** – הבנה מעמיקה של תחום זה על מגוון הסוגיות

שהוזכרו ורבות אחרות

- **טיפול יסודי בסוגיות אתיות וחברתיות** – יכולות חדשות אלה מביאות לאתגרים חסרי תקדים לחברה האנושית

Murat Durmus  
(CEO AISOMA)

## 9 ethical issues in Artificial Intelligence

### Unemployment

What happens after the end of jobs?

### Inequality

How do we distribute the wealth created by machines?

### Humanity

How do machines affect our behavior and interaction?

### Evil genies

How do we protect against unintended consequences?

### Artificial stupidity.

How can we guard against mistakes?

### Racist robots.

How do we eliminate AI bias?

### Security.

How do we keep AI safe from adversaries?

### Singularity

How do we stay in control of a complex intelligent system?

### Robot rights

How do we define the humane treatment of AI?

Content Source:  
World Economic  
Forum (WEF)

הגיע הזמן לסיים

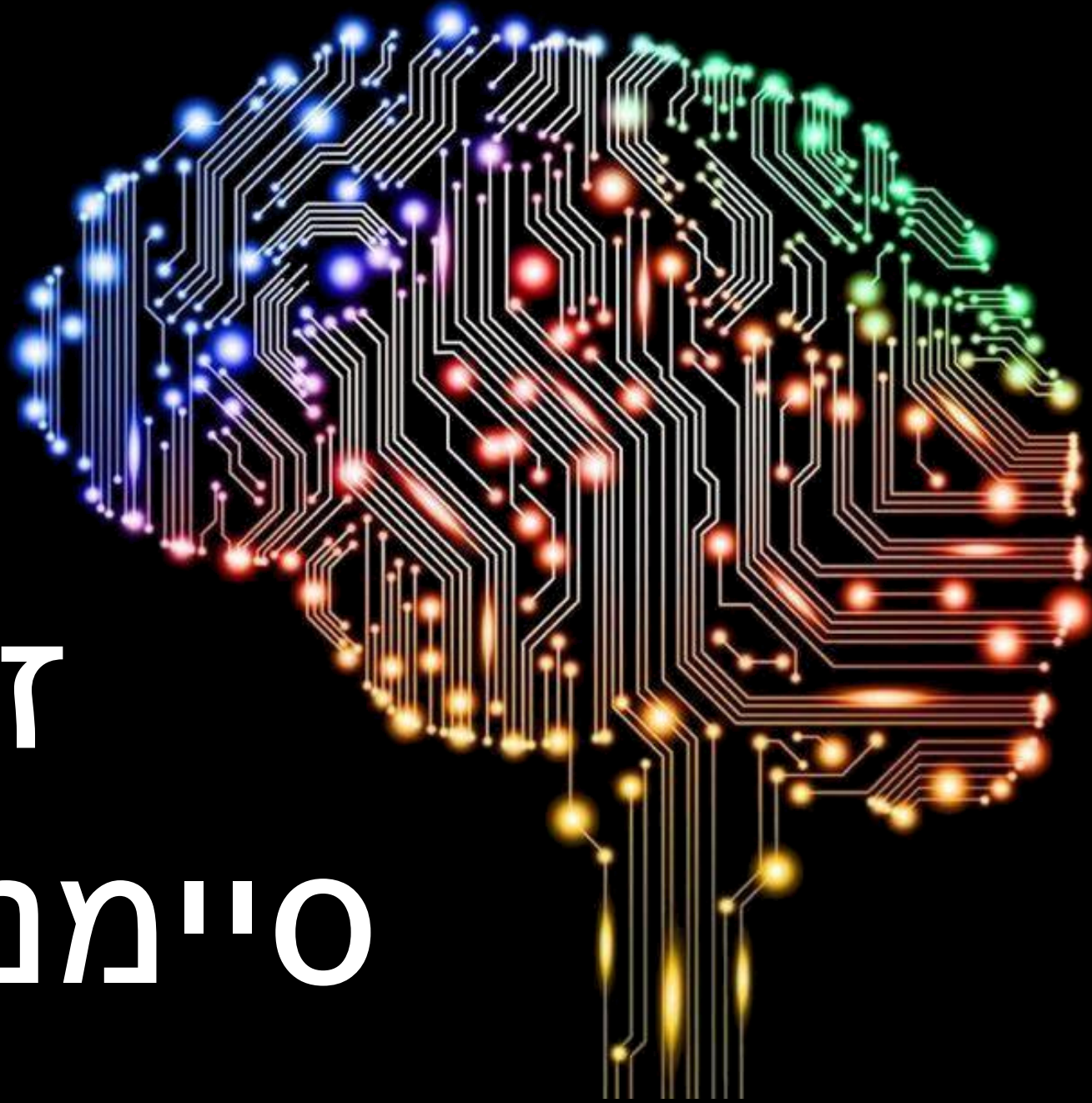
# על מה לא דיברנו ?

- ☐ על ה**טירוף** המוחלט **בכנסים** אקדמיים בשנים האחרונות בשל **מבול** של מאמרי למידה עמוקה
- ☐ על שינויים **דרמטיים** שצפויים בשוק העבודה והכלכלה בשל יכולת של מכונות להחליף בני אדם, וההשפעה על החברה
- ☐ על הביקוש ה**מטורף** באנשי למידה עמוקה בתעשייה **ומשכורותיהם** המופרכות
- ☐ איפוא **ישראל** מול העולם בזירה זו?
- ☐ איפוא ה**טכניון** מול עולם האקדמיה בזירה זו?
- ☐ על מה **אנו עובדים**?
- ☐ **מה צופן העתיד** בעולם המחקר בצל ההתפתחויות הללו?
- ☐ התמקדנו בלמידה עמוקה. מה עם **טכניקות למידה ממוחשבות** אחרות? מה עוד קיים תחת המעטה של **AI**?



ARTIFICIAL INTELLIGENCE





זהו

סיימנו !