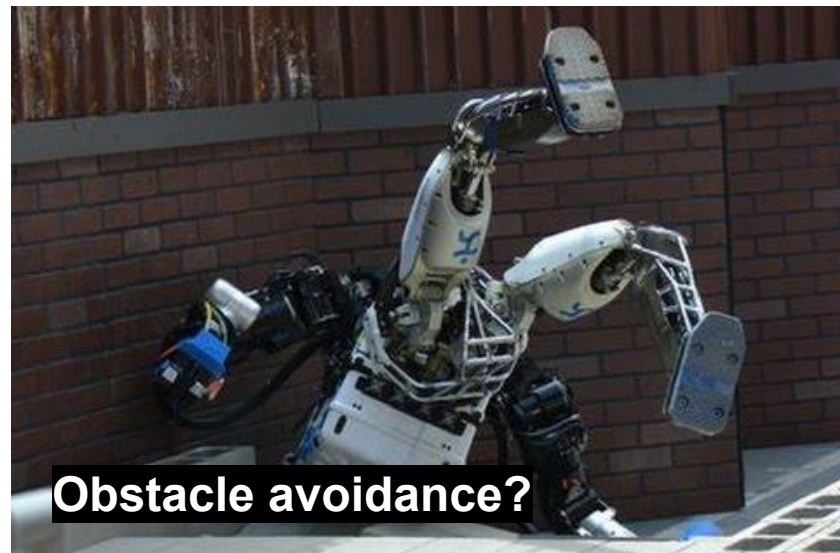
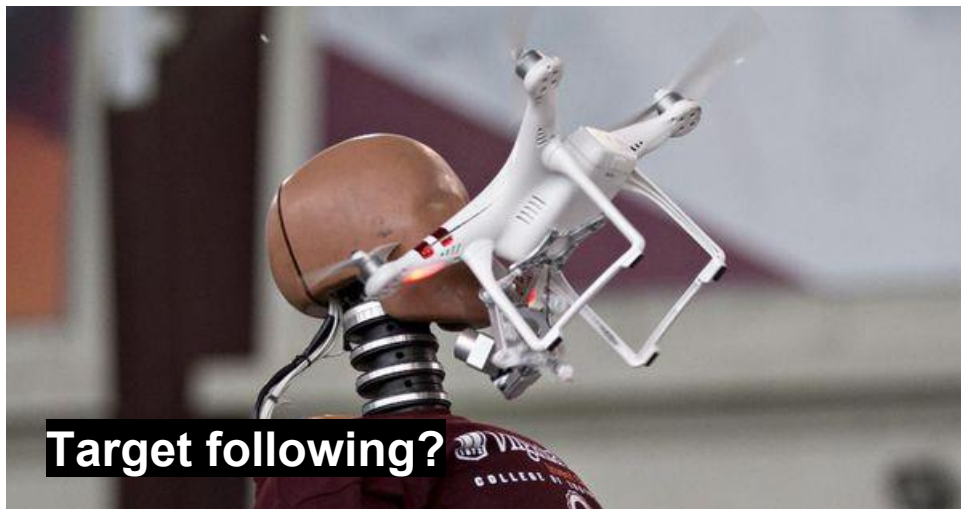


Virtual-to-Real: Learning to Control in Visual Semantic Segmentation

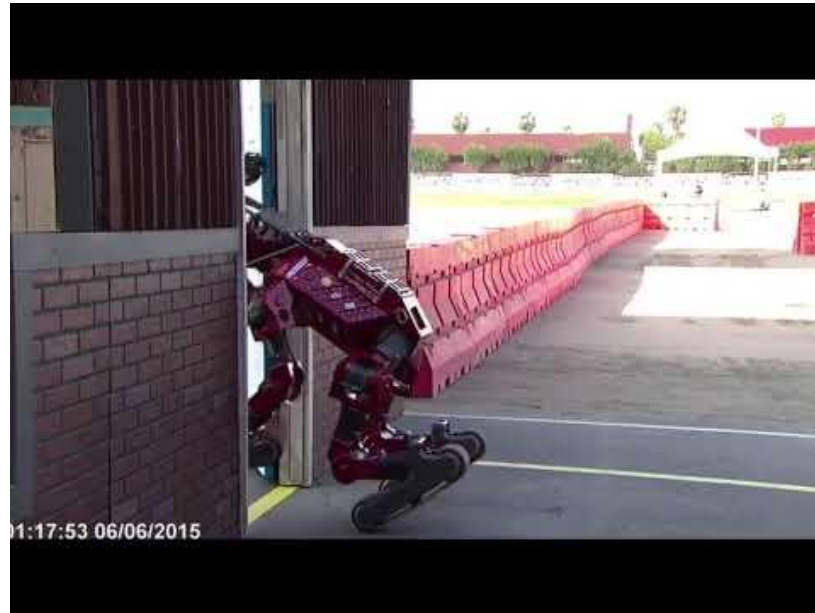


Why we need it in daily life? And how we do it?



אין צורך להזכיר מהי כמות השימוש היום יומית שאנו עושים בטכנולוגיה.
בכל שנה כמות השימוש שאנו עושים ברובוטים גדלה (שואבי אבק, רחפנים,
רכבים אוטונומיים ועוד)
אבל איך הם בעצם עובדים?
איך רכבי TESLA יודעים לזהות לבד סכנה על הכביש, לשמור על נתיב ועוד?
למעשה מדובר באלגוריתמי למידה עמוקה חכמים אשר כדי לעבוד באופן
אופטימאלי ולהיות יעילים צריכים כמויות מידע עצומות (זו הסיבה שכל רכב
טסלה למשל שולח כמות עצומה של נתונים לשרתי החברה בכל נסיעה שמבצע
הבעלים).

Training is not that easy nor a cheap task



So lets provide the data and train them, simple?

איסוף מידע מן העולם הפיזי למטרות לימוד אלגוריתמים הינו בד"כ פעולה הדורשת זמן רב ומשאבים יקרים ואף מסכנת רובוטים שבירים. התקדמויות אחרונות בתחום לימוד רובוטים תומכות בשימוש של סימולטורים כפלטפורמת אימונים. למרבה הצער, הפער בין מידע (ויזואלי) סינטטי לאמיתי אינו מאפשר השמה ישירה של מודלים אשר עברו אימון בעולם וירטואלי לאילו שאומנו בעולם האמיתי.



המטרה הינה לאמן רובוטים בסביבה וירטואלית
כך שיוכלו גם ליישם את הנלמד גם בעולם הפיזי.



So what's the big deal, right?

לעיתים קרובות, רובוט אוטונומי צריך "להבין" את הסביבה הויזואלית בה הוא נמצא על מנת לנווט ליעד ספציפי, תרגום וייצוג של קלטים ויזואליים על מנת לבצע פעולות ואינטראקציה עם אובייקטים. אולם, הדבר מאתגר לרובוטים בתוך סביבה לא ממופה, כגון תמונה צבעונית מורכבת עם "רעשים".





Existing proven approaches

קיימות שתי גישות שהראו עצמן כיעילות (modular, and-to-and learning based) במגוון משימות תנועת רובוטים המונחות ראיה. השיטות הראו הצלחה במשימות של חיפוש יעדים ספציפיים ע"י רובוטים. אומנם השיטות הללו נראות אטרקטיביות אך הן דורשות כמויות עצומות של מידע (training data) בכדי לאמן את הסוכנים, אך איסוף מידע זה הינו יקר, צורך זמן רב ומשאבי אנוש רבים.



What is the alternative approach

האלטרנטיבה המוצעת היא ע"י לימוד הרובוטים בעולם וירטואלי.
האלטרנטיבה מאפשרת זירוז של תהליך הלימוד ושיפור היעילות שלו. כמו כן, אלטרנטיבה זו מאפשרת הפחתה ניכרת בעלויות.
הבעיה שרובוט שעבר אימון בעולם וירטואלי אינו יכול פשוט ליישם את מה שלמד בעולם הפיזי עקב הבדלים ברורים בין העולמות ורמת הדיטליזציה (פרטים) ביניהם.

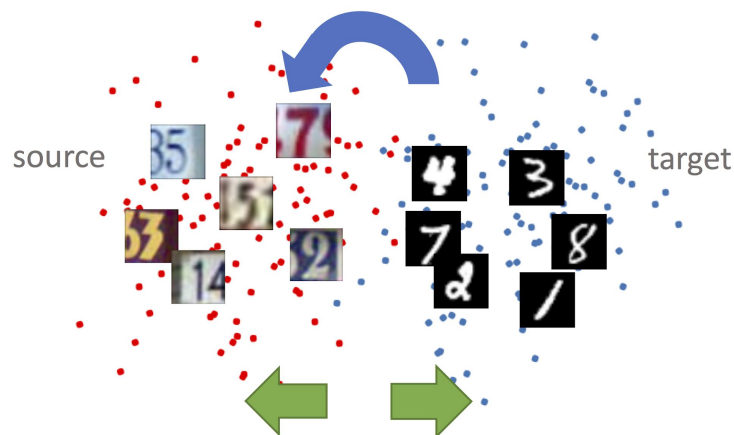


What was done so far in this area

ישנן מס' טכניקות (domain adaptation, domain randomization) שניסו לאמן רובוטים בעולם וירטואלי, האחת בעזרת סימולטורים והשניה בשימוש צורות אובייקטים שונים, צבעים, צורות תאורה שונות ועוד. הגישות אפשרו השמה של הנלמד בעולם האמיתי אך כפי שצויין קודם כמויות המידע שהיה צורך לאסוף ולתייג היו גדולות ויקרות. הגישה השניה DR אומנם לא הייתה זקוקה למידע מן העולם הפיזי אך גם היא הוכיחה עצמה כלא יעילה במיוחד עקב שלוקה בחוסר לימוד סיסטמטי שיכול לקבוע אילו פרמטרים יכולים להיות רנדומליים וכמו השיטה הקודמת גם כן זקוקה לכמות עצומה של דגימות כדי לעבוד.

נראה כעת את הגישות...

Domain Adaptation

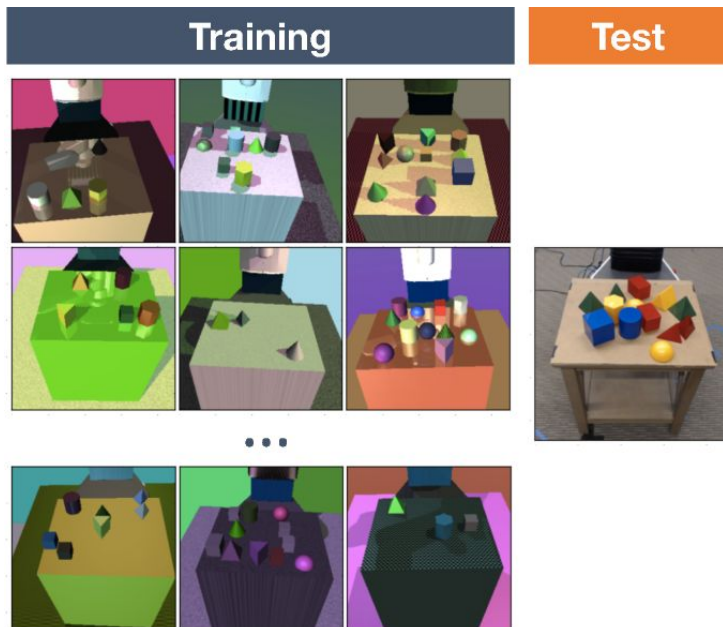


הגישה בעיקר עוסקת בשיפור מודל שכבר עבר אימון בעולם וירטואלי תוך שימוש במידע מן העולם האמיתי לצורך כיוול סופי.

זהו תחום הקשור ללמידת מכונה והעברת המידע הנלמד.

תרחיש זה מתעורר כאשר אנו שואפים ללמוד מתוך הפצת נתונים ממודל המקור אל מודל היעד תוך ביצוע הפצת נתונים שונים (אבל קשורים) לדוגמא ספרות 0,1,2,3,4,5,6,7,8,9.

Domain Randomization



A simple technique for training models on simulated images that transfer to real images by randomizing rendering in the simulator.

With enough variability in the simulator, the real world may appear to the model as just another variation.

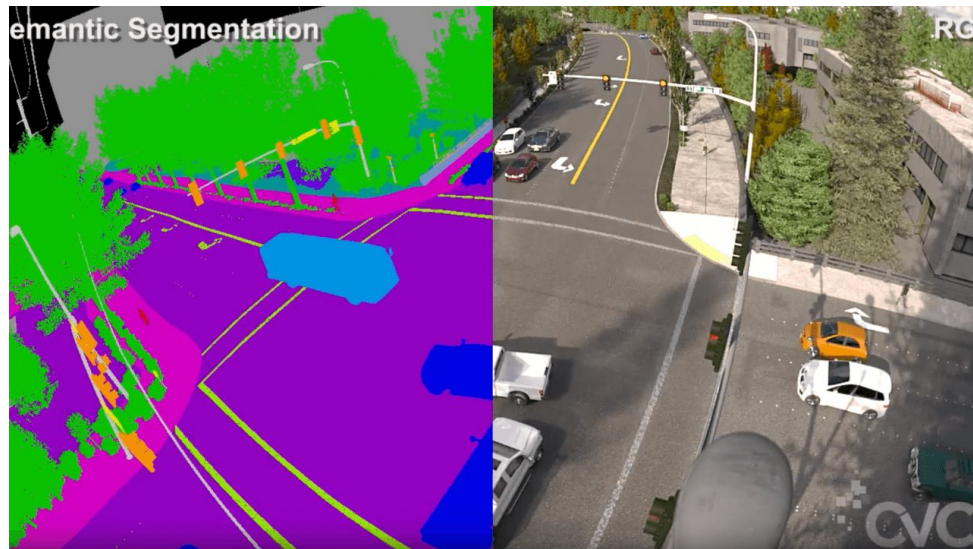
The technique **lacks** a systematic way to determine which parameters are required to be randomized.



New modular architecture

הגישה שלנו מבוססת על למידה נוירונית עמוקה (Deep Neural Network), שבאה להפריד בין **מודל התפיסה ומודל השליטה** כאשר שני המודלים עוברים אימון באופן נפרד כאשר הגורם המקשר בין המודלים הינו סגמנטציה סמנטית של תמונה. כל מודל עובר אימון בנפרד ואינו מניח שקיים מידע קודם כלשהו מן המודל השני והם בלתי תלויים.

Semantic Image Segmentation



תחילה נסביר מהי סגמנטציה של תמונה בה עושה שימוש המודל הראשון של השיטה.

הסגמנטציה חיונית למשימות של ניתוח תמונות. פילוח סמנטי מתאר את התהליך של שיוך כל פיקסל של תמונה עם תווית מחלקה (כגון פרח, אדם, כביש, שמים, אוקיינוס או רכב).

הסגמנטציה הסמנטית שומרת רק את המידע החשוב ביותר, דבר היוצר סדר ומבניות בייצוג של האובייקטים.

השמות עיקריות לפילוח סמנטי כוללות: נהיגה אוטונומית, הימנעות ממכשולים בזמן טיסת רחפן ועוד.

Segmentation example

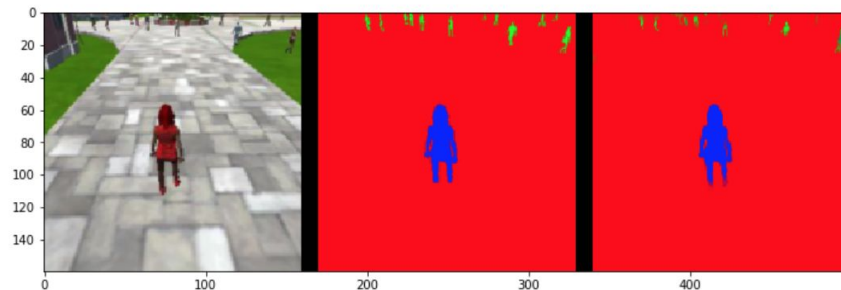


(Start at 0:55)

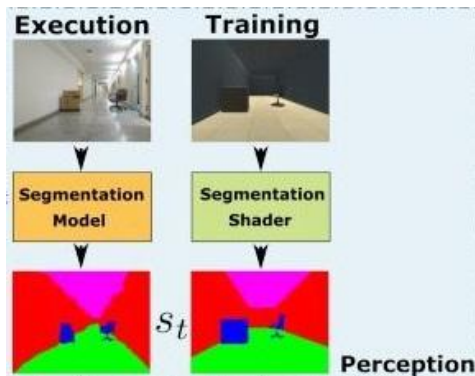
Courtesy of IEEE and IJCV

1# Perception module

מודל התפיסה מתרגם תמונות RGB למקטעים סמנטיים, הרעיון הוא שניתן לתרגם גם תמונות מן העולם הפיזי וגם מן העולם הוירטואלי בצורה דומה. אימון וכיול של מודל התפיסה לצורך סגמנטציית תרחישי חוץ (כבישים, אזורים פתוחים וכדומה) כבר הפך היום לפשוט.



Perception module



המטרה העיקרית של מודל התפיסה הינה לייצר סגמנטציה סמנטית של תמונה ולהעבירה למודול השליטה. בשלב האימון מתבצע עיבוד תמונה הנלקחת מסצינה של העולם הוירטואלי. לאחר מכן, בשלב היישום מוזנת תמונה של העולם האמיתי תוך שהיא עוברת אנליזה של מס' אלגוריתמי למידה עמוקה אשר עברו אימון לתרחישי חוץ ופנים. למרות שהתמונות מן העולם האמיתי שונות (ברמת הפרטים ורעשים) מתמונות מן העולם הוירטואלי, הסגמנטציה הסמנטית המתקבלת כמעט זהה. עובדה זו היא המאפשרת לסגמנטציה הסמנטית להיות הגשר המקשר מן העולם הוירטואלי לאמיתי.

The hardware price is lower

בנוסף, לצורך סגמנטציה סמנטית (כדי ללכוד את התמונה) נדרשת בסך הכל מצלמה מונוקולרית פשוטה לעומת המכשור היקר כגון מצלמות עומק ומאתרי טווח מבוססים לייזר (מכשירים אילו יכולים לסייע בסגמנטציה של התמונה ברמת החומרה).

Monocular camera



VS

Laser rangefinders

• Internal Display

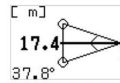


Measurement example (Three-point measurement: height between two points)

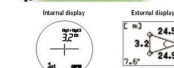
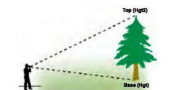


When the measurement is completed, the height between points 1 and 2 is displayed on the internal LCD as the HgtHgt2 value, and Hgt1 and Avg1 are shown on the external LCD. Points 1 and 2 are indicated.

• External Display



Measurement example (2-point height measurement)



When the measurement is completed, you see the height from the base to the top displayed on the internal LCD with Hgt + Hgt2. For more information, refer to the manual.



Depth camera



It's faster and more efficient approach

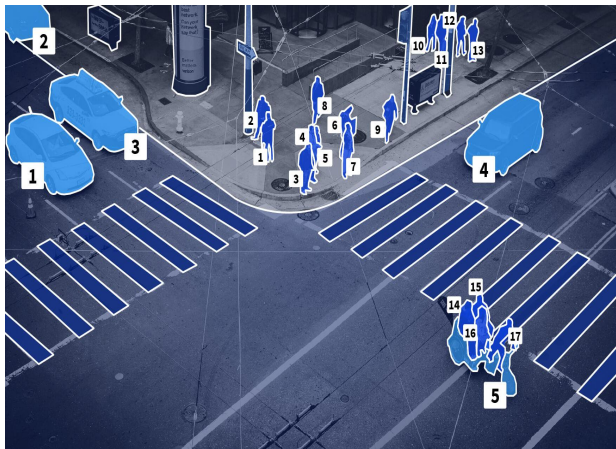


הארכיטקטורה הנ"ל מספקת יעילות טובה יותר מן השיטות הקונבנציונליות כיוון שסימולטורים (הדמיות) יכולים לבצע עיבוד מהיר יותר של מקטעי תמונה ממופה מאשר תמונה רגילה המוצפת לרוב בפרטים לא רלוונטיים ורעשים אחרים.

יתרון חשוב נוסף של השיטה הינו המודולריות שלה כי ניתן להחליף כל אחד מן השלבים במודל אחר (נרחיב בהמשך).

נראה דוגמא לסגמנטציה סמנטית בזמן אמת...

2# Control policy module



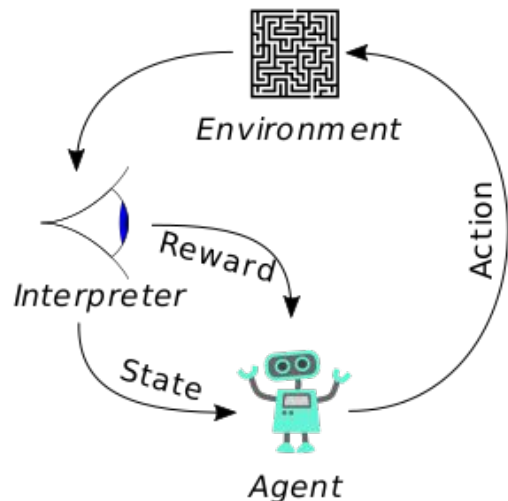
זוהי דוגמא לקלט של מודל התפיסה

מודל השליטה הינו אחראי על מציאת מדיניות שליטה אופטימלית פר משימה מסוימת. בשלב האימון של המודל נעשה שימוש בסביבות וירטואליות בלבד.

מודל השליטה מיישם מטודות של למידת חיזוק (reinforcement learning) ומקבל בתור קלט מקטעי תמונה סמנטיים, אילו יכולות להיות תמונות שנוצרו על ידי סימולטורים או מקטעי תמונות ממאגרים שעברו עיבוד קודם.

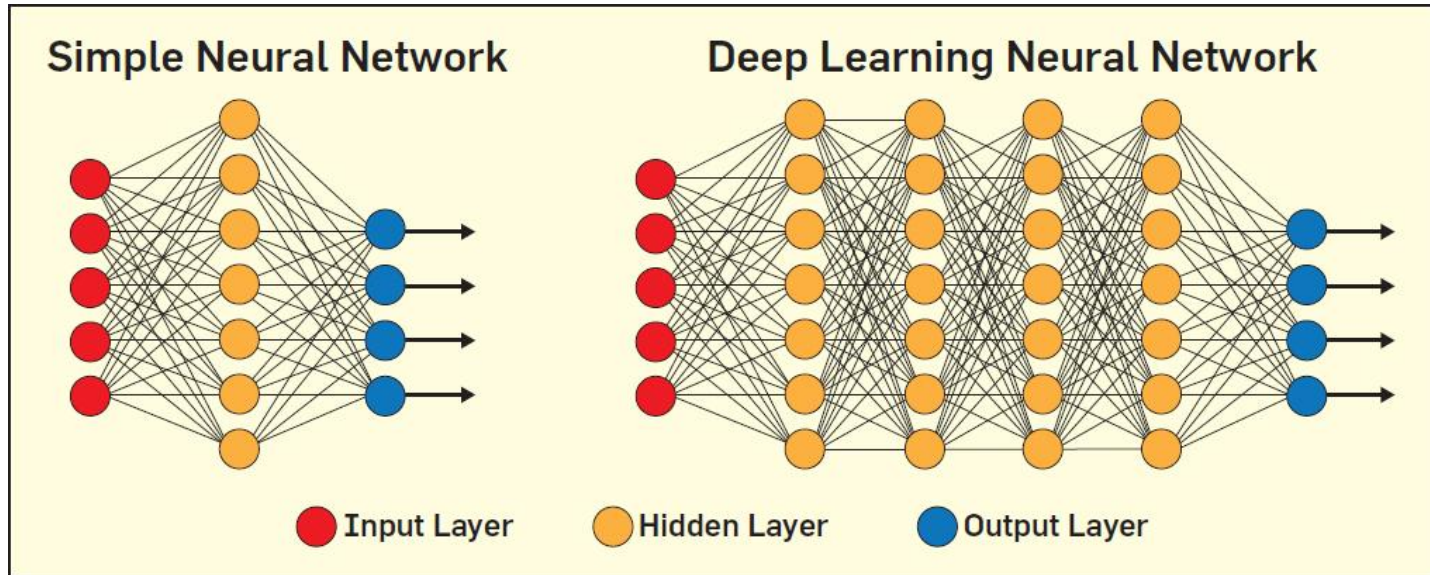
סוכן של למידת חיזוק נמצא באינטראקציה עם הסביבות הוירטואליות ואוסף מידע (training data). ובשלב הביצוע מודל השליטה מקבל מקטעי תמונות סמנטיים ממודל התפיסה. תמונות אלו עברו עיבוד בשלב הראשוני במודל התפיסה ודבר זה מאפשר לסוכן אינטראקציה עם העולם הפיזי.

What is Reinforcement Learning



למידה באמצעות חיזוקים, היא שיטת למידה המבוססת על ביצוע פעולות אשר נסמכות על חיזוקים חיוביים או שליליים. מודל הלמידה באמצעות חיזוקים מבוסס על למידה מתוך ניסוי וטעיה שמתבצעת בסדרת אינטראקציות בין סוכן לומד (רכיב תוכנה או רובוט) לבין סביבתו המיוצגת כמידע חושי (סנסורי) הזמין לסוכן הלומד. בכל אינטראקציה כזאת, מקבל הסוכן מידע מסביבתו לגבי המצב הנתון ובוחר לבצע פעולה ממרחב הפעולות הזמין לו בהתאם למדיניות הפעולה הנוכחית. בתגובה מתקבל אות תגמול שהוא ערך מספרי המבטא את חיוניות הפעולה בהתאם למצב ועל פיו מעדכן הסוכן את מדיניותו עבור האינטראקציות הבאות. מטרתו של הסוכן הלומד היא לגבש מדיניות פעולה הממפה בין מרחב המצבים האפשריים למרחב הפעולות כך שהתגמול הכולל (המבטא כסכום משוקלל של אותות התגמול שהתקבלו) עבור ביצוע פעולות על סמך מדיניות זו יהיה גבוה ככל שניתן.

Deep Neural Network model





Virtual-to-Real Evaluation Setup: Simulator

בשלב זה נציג את שלבי האיבולוציה של הארכיטקטורה המוצעת גם בסביבה הוירטואלית וגם בעולם האמיתי.

נתחיל בסימולטור, לצורך פיתוח סביבה וירטואלית השתמשנו במנוע Unity-3D כדי לאמן את סוכני למידת החיזוק (Reinforcement Learning). יכולת המנוע Unity-3D מאפשרת התנהגות מציאותית לאובייקטים שנוצרו בו, דבר התורם רבות לדיוק הלמידה. בעולם וירטואלי זה, סוכני למידת החיזוק מקבלים את הקלטים בצורה סגמנטציה סמנטית בגוף ראשון. לאחר מכן הסוכנים מחליטים על פעולות אשר יבצעו בהתאם לקלט. הסביבה הוירטואלית מגיבה בהתאם לפעולות הסוכנים, ותגובתה של הסביבה מתורגמת בתור אות חיזוק אצל כל סוכן. הצורות הגיאומטריות והמימדים של כל סוכן הותאמו במיוחד להיות זהים לזה של רובוט אמיתי.



Virtual-to-Real Evaluation Setup: Tasks

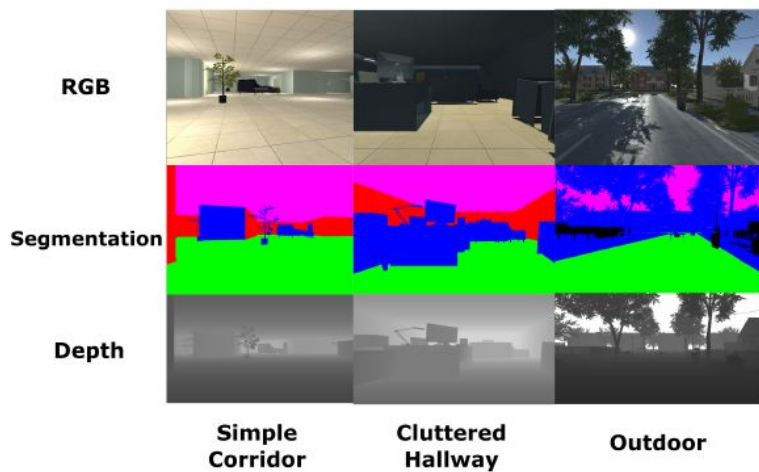
מודל זה מוצע כנגד המודלים המתחרים בסביבות וירטואליות ואמיתיות על בסיס השוואת הצלחתם בשני משימות בסיסיות שכמעט כל רובוט חייב ליישם: obstacle avoidance, target following.

בתחילת שני המשימות, סוכני למידת החיזוק התחילו במיקומים רנדומליים בסביבות בהן הוצבו. בכל מקטע זמן, הסוכן היה צריך לבחור פעולה מבין שלושת הפעולות: לנוע קדימה, לפנות שמאלה, לפנות ימינה.

מטרת הסוכן במשימת **Obstacle Avoidance** הינה ניווט ללא התנגשויות בסביבה מגוונת עם תרחישים שונים (למשל עץ שקורס מולו).

מטרת הסוכן במשימת **Target Following** הינה לנוע אחר מטרה זוה תוך כדי הימנעות ממכשולים.

Virtual-to-Real Evaluation Setup: Scenarios



נעריך את המודל בעזרת שלושת התרחישים הבאים:

1. **מסדרון פשוט:** התרחיש משלב בתוכו מעברים פנימיים ישרים, פניות חדות, מכשולים סטטיים (כגון שולחנות, כסאות, קופסאות וכו') וגם מכשולים נעים (אנשים הולכים).
2. **מסדרון עמוס:** התרחיש משלב מסדרון עמוס במכשולים סטטיים ומכשולים זזים לצורך הערכת יכולת הסוכן בתימרון יעיל והימנעות ממכשולים אילו בסביבה צפופה.
3. **באיזור פתוח (חוץ):** תרחיש המשלב כביש עם מדרכות, בניינים, מכוניות בתנועה והולכי רגל.
בחירת התרחישים הנ"ל נעשתה לצורך הערכת מידת הדיוק ומידת יעילות האימון שעבר מודל השליטה.



Experimental Results & Conclusion

הארכיטקטורה החדשה הוכיחה את עצמה במבחן התוצאה בכל התרחישים השונים בה נבדקה אל מול המודלים הקיימים עם תוצאות טובות יותר. הרעיון העיקרי שלה מבוסס על הפרדה לשני מודלים (מודל התפיסה ומודל השליטה) תוך שימוש בקונספט של סגמנטציה סמנטית של תמונה המקשר בין שני המודלים. המודל עבר אימון בעזרת למידת חיזוקים (Reinforcement Learning) וללא שימוש במודל Domain Randomization שכבר ציינו את חסרונותיו. הבדיקות שנעשו כללו משימות- obstacle avoidance, target following שהציגו בתוצאתם עליונות על המודלים הקיימים בשני העולמות (וירטואלי ואמיתי). כמו כן המודל הוכיח את המודולריות שלו במעבר בין שני העולמות ללא כוילים נוספים.

Example in real life applications

דוגמא גם למעקב אחר יעד
ספציפי תוך כדי הימנעות
ממכשולי המסלול

הדגמה של רחפן Intel



Summary



סיכום התרומות של השיטה:

א. פתרון קל למימוש ואפקטיבי לבעיה מורכבת (מוירטואלי לפיזי).

ב. ארכיטקטורה מודולרית חדשנית המבוססת על שני מודלים, התפיסה והשליטה.

ג. קונספט חדשני המגשר על הפער בין הפיזי לוירטואלי בשימוש של מקטעי תמונה ממופים.

ד. שיטה להחלפת סביבת פעולה של רובוט ללא כוילים\התאמות של מודול מדיניות השליטה.

ה. מודל הנחיה חזותית המשנה את התנהגותו של רובוט תוך הסתגלות לשינויים בסביבתו.

Thank you

Zhang-Wei Hong, Yu-Ming Chen,
Hsuan-Kung Yang, Shih-Yang Su,
Tzu-Yun Shann, Yi-Hsiang Chang, Brian
Hsi-Lin Ho, Chih-Chieh Tu, Tsu-Ching
Hsiao, Hsin-Wei Hsiao, Sih-Pin Lai,
Yueh-Chuan Chang, Chun-Yi Lee

Elsa Lab, Department of Computer
Science, National Tsing Hua University.

