

Virtual-to-Real: Learning to Control in Visual Semantic Segmentation
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, Hsinchu,
Taiwan
<https://www.ijcai.org/proceedings/2018/682>

מדיגיטלי למוחשי – לימוד שליטה

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.

מאמר זה מציע ארכיטקטורה מודולרית להתמודדות עם בעיית המעבר מדיגיטלי למוחשי.

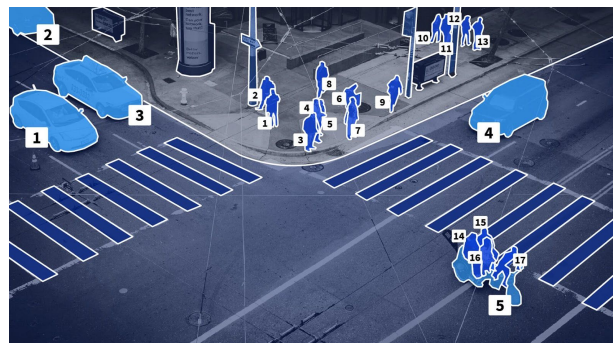
הארכיטקטורה המוצעת מפרידה את מודל הלמידה לשני מודולים: תפיסה, ומדיניות בקרה ומשתמשת בסגמנטציה סמנטית (של תמונה) לשם מטא-ייצוג על מנת לקשר בין שני המודלים.

מודול התפיסה מתרגם כל תמונת RGB אשר נקלטת לייצוג סמנטי של מקטעים. Control policy module מיושם בתור תהליך למידה עמוקה, שמבצעת פעולות המתבססות על סגמנטים שתורגמו מן התמונה. הארכיטקטורה שלנו מחושבת על בסיס obstacle avoidance (הימנעות ממכשולים) וגם target following (מעקב אחר מטרה). תוצאות הניסויים הראו שהארכיטקטורה שלנו משפרת באופן משמעותי את תוצאות של כל המודלים האחרים, גם בסביבות וירטואליות וגם בסביבות אמיתיות, ומדגימה עקומת למידה מהירה יותר מהם.

שליטה על מערכות רובוטיות המבוססות על תפיסה חזותית מושכת תשומת לב בשנים האחרונות. המערכות הרובוטיות מתנהגות כקלטים ויזואליים המכילים מידע עשיר על העולם הפיזי. לעיתים קרובות, רובוט אוטונומי צריך "להבין" את הסצנה 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". זהו דבר קשה במיוחד לתכנן רובוט שכזה, העונה על הדרישות הנ"ל.

שני הגישות (modular, and-to-and learning based) הוכיחו עצמן כיעילות במגוון משימות המבוססות "ראיה" (קלט של תמונה). לדוגמא ניווט רחפן על קטע שביל שתפקידו להתמקד על דמות בן אדם ולעקוב אחריה במהלך הטיסה תוך הימנעות ממכשולים שונים כגון עצים, אומנם גישות כאלו המבוססות למידה מאוד "חמות" היום בשוק, הן דורשות כמויות מידע "לימודי" עצומות כדי להיות יעילות, ומידע זה גם צריך צריך להיות ממופה נכון וזו משימה יקרה הדורשת משאבים רבים.

האלטרנטיבה שהמאמר מציע לזירוז תהליך הלימוד ושיפור היעילות שלו היא ע"י לימוד הרובוטים בעולם וירטואלי, הבעיה שרובוט שעבר אימון בעולם וירטואלי אינו יכול פשוט ליישם את מה שלמד בעולם הפיזי עקב הבדלים ברורים בין העולמות ורמת הדיטליזציה בינם. ישנן מס' טכניקות (domain adaptation, domain randomization) שכבר מאמנות רובוטים, 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 ולתתג 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. הגישה היעילה יותר בא דן המאמר מבוססת על למידה נירונית עמוקה DNN, שבאה להפריד בין מודל התפיסה ומודל השליטה כאשר שני המודלים עוברים אימון באופן נפרד ולא תלוי אחד בשני. כל מודל אינו מניח שקיים מידע כלשהו מן המודל השני. מודל התפיסה מתרגם תמונות GRB למקטעים סמנטיים:



מודל השליטה מיישם מטודות של למידה המקבל כקלט מקטעי תמונה סמנטיים, אילו יכולות להיות תמונות שנוצרו על ידי סימולטורים או מקטעי תמונות ממאגרים שעברו עיבוד קודם. הסוכן במודל השליטה נמצא באינטראקציה עם הסביבות הוירטואליות ואוסף מידע (training data) בתהליך. ובשלב הביצוע מודל השליטה מקבל מקטעי תמונות ממודל התפיסה המאפשר לסוכן אינטראקציה עם העולם הפיזי. בגלל שהתמונות הינן של העולם הפיזי אך עברו עיבוד בשלב הראשון לסוג של עולם וירטואלי כך הסוכן בעצם יכול להיות באינטראקציה "ישירה" (דרך שני שלבים) אם העולם הפיזי. הארכיטקטורה הנל מספקת יעילות טובה יותר מן השיטות הקונבנציונליות כיוון שסימולטורים 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 בצורה פשוטה יחסית (API) כל עוד הפורמטים של מקטעי התמונה זהים.

כדי להוכיח את יעילות השיטה נשים אותה בשני מבחנים (הימנעות ממכשולים ומעקב אחר מטרה) מול השיטות האחרות. אילו שני פעולות בסיסיות הנדרשות מרובוט אוטונומי, אחרת הרובוט לא יוכל לנוע במסלול המיועד או לחילופין יתנגש בכל מכשול בדרכו ליעד. והתוצאות כל המבחנים הוכיחו את עליונות הארכיטקטורה, מעבר לכך גם שהמודל מאפשר לימוד מהיר יותר מן השיטות האחרות.

סיכום התרומות של השיטה:

- א. פתרון קל למימוש ואפקטיבי לבעיה מורכבת (מורטואלי 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 כיולים\התאמות של מודול מדיניות השליטה.
- ה. מודל הנחיה חזותית המשנה את התנהגותו של רובוט תוך הסתגלות לשינויים בסביבתו.