|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| C:\Users\Student\Downloads\Logo_Eng_TAU_New.jpg | | | | |
| אוניברסיטת תל-אביב | |  | | **Tel-Aviv University** |
| הפקולטה להנדסה  בי"ס להנדסת חשמל | |  | | Faculty of Engineering  School of Electrical Engineering |
| זיהוי ועקיבה אחר רחפן | | | | |
| פרויקט מס' 16-1-2-1091  דו"ח סיכום | | | | |
| מבצעים: | | | | |
|  | רפאל בוזגלו | | 305081002 | |
|  | שמואל סעדי | | 205883945 | |
| מנחה: | | | | |
|  | מר רועי אורפייג | | מל"מ, תעשייה אווירית | |

**תוכן עניינים**

[**1.** **תקציר** 3](#_Toc485745760)

[**2.** **הקדמה** 4](#_Toc485745761)

[**3.** **רקע תיאורטי** 5](#_Toc485745762)

[3.1. בעיית העקיבה הכללית 5](#_Toc485745763)

[3.2. עקיבה אחר ריבוי מטרות (Multi Target Tracking). 5](#_Toc485745764)

[3.3. Harris Corner Detector 6](#_Toc485745765)

[3.3. רגיסטרציה 7](#_Toc485745766)

[3.4. החסרת תמונות בזמן 9](#_Toc485745767)

[3.5. פעולות מורפולוגיות על תמונות 10](#_Toc485745768)

[3.6. Speeded up robust features (SURF) 12](#_Toc485745769)

[3.7. רשתות נוירונים מלאכותיות 12](#_Toc485745770)

[3.8. Alex-Net 14](#_Toc485745771)

[3.9. מודל SVM 16](#_Toc485745772)

[**4.** **תכנון ומימוש** 18](#_Toc485745773)

[4.1. דיאגרמת בלוקים 19](#_Toc485745774)

[4.2. זיהוי ואפיון נקודות עניין 20](#_Toc485745775)

[4.2.1. Harris Corner Detector 20](#_Toc485745776)

[4.2.2. חיסור תמונות בזמן 21](#_Toc485745777)

[4.2.3. חילוץ מתארים לאפיון 22](#_Toc485745778)

[4.3. Multi Target Tracking 22](#_Toc485745779)

[4.3.1. שיוך נקודות תואמות 23](#_Toc485745780)

[4.3.2. עדכון מבנה הנתונים 23](#_Toc485745781)

[4.4. סיווג מטרות 24](#_Toc485745782)

[4.4.1. חילוץ מאפיינים באמצעות CNN 25](#_Toc485745783)

[4.4.2. סיווג באמצעות מודל SVM 26](#_Toc485745784)

[**5.** **ניתוח סיכום והצעות להמשך** 28](#_Toc485745785)

[5.1. חשיבות שלב הסיווג לתהליך העקיבה 28](#_Toc485745786)

[5.2. ניתוח תוצאות 29](#_Toc485745787)

[5.3. סיכום 30](#_Toc485745788)

[5.4. הצעות להמשך 30](#_Toc485745789)

[**6.** **ביבליוגרפיה** 30](#_Toc485745790)

[**7.** **רשימת איורים וטבלאות** 31](#_Toc485745791)

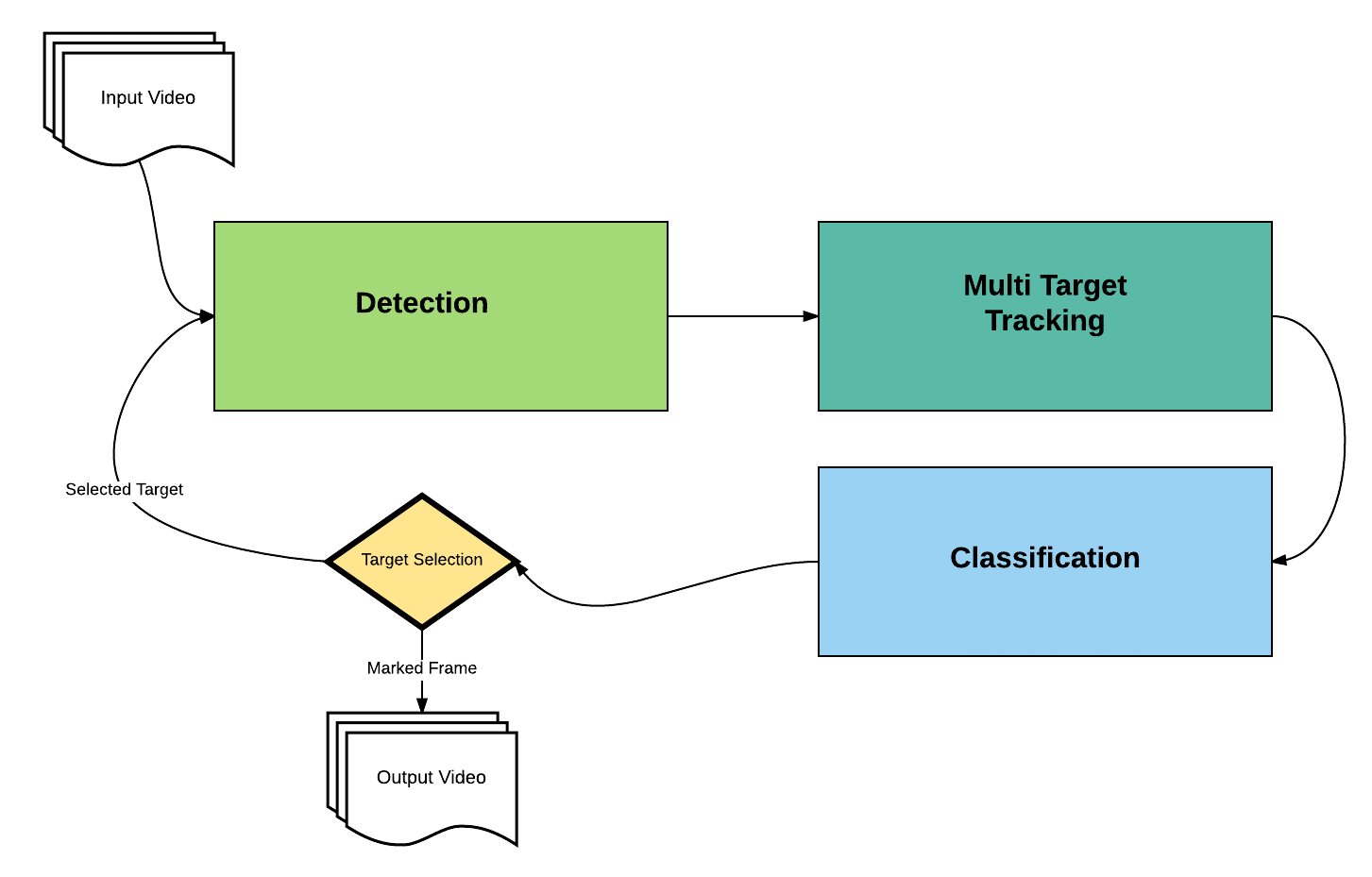
# **תקציר**

בפרויקט זה מימשנו אלגוריתם שמסוגל לבצע עקיבה אחר רחפן.

כקלט מקבל האלגוריתם וידאו, אשר אנו מניחים כי קיים בו רחפן. לאורך כל שלבי האלגוריתם, העבודה מתבצעת על כל פריים בצורה סדרתית. תוך התחשבות במאפייני הרחפן, נחלץ תחילה נקודות עניין במספר שיטות המתבססות על עקרונות שונים במטרה לכלול את הרחפן בקבוצת נקודות אילו. בכך נצמצם את היקף הבעיה, מחיפוש בפריים כולו, למספר אזורים בולטים.

על מנת שנוכל להתאים נקודות אילו לאובייקטים הקיימים, נתאר אותן באמצעות סביבתן, ובכך נוכל לייצר עקיבה רציפה ולאורך זמן. כעת, כדי להבטיח כי האובייקט אחריו אנו עוקבים אכן רחפן, נעזר במסווג ייחודי אותו יצרנו אשר יודע לזהות רחפנים.

לאחר רכישת המטרה נבצע שלבים אילו על הפרימיים הבאים תוך התחשבות בממצאי האלגוריתם עד כה.



איור 1 : דיאגרמת בלוקים כללית

# **הקדמה**

בעולם המתפתח טכנולוגית שימוש בכלי תעופה קומפקטיים אשר מאוישים מרחוק הולך וגובר. כלי תעופה אילו זולים, נגישים לכל דורש והשימוש בהם אינו מצריך רישיון הטסה ואכיפה בנושאים אילו עדיין לא הבשילה לכדי מעשה.

הגחתם של הרחפנים לכלל הציבור יוצרת אתגרים רבים, הבולט מבניהם הוא האתגר הביטחוני אשר איתו נאלצים גופים ומדינות רבות להתמודד כבר בימים אילו. האיומים כוללים איסוף מודיעין, העברה בלתי חוקית של סחורה ללא פיקוח, ואף תקיפה מטרות. בשל גודלו מהווה הרחפן כלי אידאלי לביצוע פעולות אילו.

כיום, כדי להתמודד עם בעיות דומות, גילוי כלי תעופה, נעשה שימוש במכ"ם, מערכת יקרה, לא נגישה ומיועדת לזיהוי עצמים גדול יותר (מרחפן). לאור עובדות אילו, ניתן לראות מגמה של שימוש באלגוריתמי ראייה ממוחשבת לפתרון בעיות דומות.

הפרויקט בוצע במסגרת התעשייה האווירית. קיבלנו לידנו סרטונים בהם צולם רחפן ממרחקים שונים ומזוויות שונות. מטרתנו הייתה לתכנן אלגוריתם לגילוי הרחפן, לבצע עקיבה אחריו ובהמשך גם לזהות שאכן מדובר ברחפן. האלגוריתם שפיתנו הניב תוצאות מרשימות והוא סייע לקידום מערכת חדשנית לאיתור ויירוט רחפני אויב.



איור 2: תמונת כניסה לאלגוריתם הכילה רחפן אותו קשה לזהות. מימין בעיגול גילוי האובייקט באמצעות האלגוריתם.

בתחומי הגילוי, עקיבה וזיהוי פותחו אלגוריתמים רבים המתבססים על הנחות כגון, שינויים איטיים בזמן, רקעים סטטיים בקירוב, היסטוגרמת הצבעים של האובייקט משתנה בצורה איטית, וכן גודלו וכיוון תנועתו. ההנחות אילו אינן תקפות במלואן במקרה של עקיבה אחר אובייקט קטן מאוד כגון רחפן. התבססות לא מושכלת על הנחות אילו תניב תוצאות ירודות.

בשל כך, גישת הפתרון שלנו כוללת שימוש משולב באלגוריתמים מתחומים שונים אשר התאמנו אותם לצורך זה. לדוגמא שילוב של אלגוריתם עקיבה אחר ריבוי מטרות (multi target tracking), והתבססות על מאפייני מרחביים בתמונה (Descriptors) הן במרחב והן בזמן. בנוסף שילבנו באלגוריתם מסווגים מתחום הבינה המלאכותית (deep learning) - השתמשנו ברשת נוירונים ובנינו מסווג (SVM) אשר במהלך העקיבה בודק באופן רציף האם האובייקט הנעקב הוא אכן האובייקט המיועד (רחפן) ומסייע במקרים בהם אובייקט אחר חותך את הרחפן (כגון ציפור, רכב ברקע וכו').

פרק 3 יכלול רקע ספרותי אודות השיטות בהם השתמשנו בפרויקט, בפרק 4 נתאר את האלגוריתם שפותח ובפרק 5 נתאר ביצועים ותוצאות.

# **רקע תיאורטי**

# בעיית העקיבה הכללית

בבעיית העקיבה, אנו מבקשים למצוא אובייקט מסוים לאורך סרטון וידיאו שלם באופן רציף, כלומר לדעת את המיקום שלו בכל פריים ופריים.

ישנן מספר גישות שכיחות בהן משתמשים לשם עקיבה:

**Dense Optical Flow** – בגישה זו, משערכים עבור כל פיקסל ופיקסל את וקטור התנועה שלו בכל פריים.

**Sparse Optical Flow** – בגישה זו עוקבים אחר מספר נקודות עניין בתמונה, בניגוד לגישה הקודמת בה משערכים וקטור תנועה לכל פיקסל, אך במקרה זה יש כעת צורך בנקודות עניין טובות.

**Density Function Estimation** – בגישה זו משערכים את פונקציית צפיפות ההסתברות של האובייקט (של היסטוגרמת הצבעים שלו כמו ב CamShift) ומנסים לאתר את החלון בעל ההסתברות המקסימלית.

בעקיבה אנו לא מזהים את האובייקט אלא משערכים את המיקום שלו לפי מאפיינים שאנו אוספים עליו (כל גישה אוספת מאפיינים אחרים). עקיבה באופן כללי היא הרבה יותר מהירה מזיהוי האובייקט, וכן יותר רובסטית ויכולה להתמודד עם הסתרות, ושינויים בנראות האובייקט.

# עקיבה אחר ריבוי מטרות (Multi Target Tracking).

Multi target tracking (MTT) הינה סוגיה מרכזית בתחום העקיבה, המתרכזת בשערוך המיקום של מספר לא ידוע של מטרות נעות. היום, מימושי MTT שונים יכולים להימצא במגוון רחב של שימושים כגון מצלמות מעקב, ביטחון גבולות יבשתיים ואוויריים, רובוטיקה וכן במערכות רכבים אוטונומיים.

המידע המתקבל כקלט ל MTT יכול להכיל מידע רלוונטי, ממטרות אמתיות, או מידע שגוי המגיע מהתרעות שווא (clutters). ניתוח נכון של המידע והכרת התכונות הסטטיסטיות של האובייקטים והרעשים, מאפשרת את סינון התרעות השווא והתמקדות במטרות האמתיות.

הבעיה המרכזית בעקיבה אחר מספר אובייקטים, בניגוד לבעיית העקיבה אחר אובייקט בודד, היא שיוך המדידות הנוכחיות למטרות הידועות עד כה (ומציאת המדידות שלא תואמות לאף מטרה ידועה). בעיית שיוך מדידות למטרות הינה בעיית שיוך מידע מוכרת ((data association problem.

באופן ספציפי יותר, עבור המקרה של עקיבה בווידאו, המטרה היא למצוא את ההיפותזה , המקשרת בין המדידות בפריים , למטרות הידועות את כה, כלומר עד לפריים *. יתר על כן ניתן לתייג את כל המטרות למחלקות שונות ובכך למקד את ההתייחסות לכל מטרה על פי מאפייני המחלקה אליה היא משתייכת.*

# Harris Corner Detector

מטרת האלגוריתם הינה מציאת כל הפינות בתמונה נתונה. מכיוון שפינות מגלמות בתוכן שינוי עוצמתי של הגרדיאנט (בשני כיוונים) בתמונה, נרצה לחפש נקודות אשר מקיימות דרישה זו.

בהינתן תמונת גווני אפור , נעביר חלון (כאשר התזוזה בכיוון הינה בשיעור ובכיוון בשיעור ) על התמונה ונחשב את שינוי של רמות האפור בה:

כאשר:

* הינו חלון הממוקם בנקודה .
* הינה רמת האפור בנקודה .
* הינה רמת האפור בנקודה .

מאחר ואנו מחפשים את החלונות בהם מופיעות הפינות, אנו מחפשים את החלונות עם שינוי גווני אפור גדול, ולכן נרצה למצוא את הנקודות אשר ממקסמות את הביטוי הנ"ל, ובפרט את הביטוי:

נפשט את הביטוי תוך שימוש בקירוב טיילור:

נפתח סוגריים ונצמצם בהתאם:

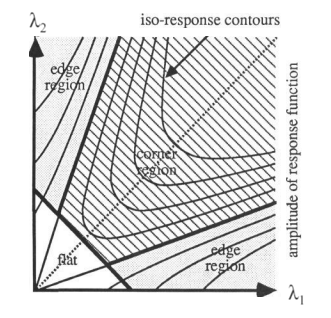
ובצורה מטריציונית:

לצורך נוחות נסמן:

ונקבל את הנוסחה:

באמצעות ניתוח הערכים העצמיים של ,, נוכל להאפיין את השינוי ברמות האפור ע"י הקשר הבא:

1. אם שני הערכים העצמיים נמוכים () אזי הפיקסל הנתון לא מהווה נקודת עיניין (איזור חלק).
2. אם אחד הערכים העצמיים נמוך ואילו השני גבוה () אזי הפיקסל הנתון מהווה שפה.
3. אם שני הערכים העצמיים גבוהים () אזי הפיקסל הנתון הינו פינה.



איור 3: גרף המתאר את איפיון הפיקסל אל מול הערכים העצמיים

האריס וסטפנס במאמרם [1] הבחינו כי חישוב מדויק של הערכים העצמיים הינו תהליך חישובי יקר, ועל כן הציעו ביטוי חלופי להערכת עומצם של הערכים העצמיים של המטריצה :

כאשר הינו קבוע רגישות (ערכיו האופייניים 0.04-0.15).

נסביר את הקשר בין הערכים העצמיים לביטוי . עבור הינו ערך גבוה () זיהינו פינה, כאשר הינו ערך שלילי גבוה בערכו המוחלט (( זוהי שפה וכאשר הינו ערך נמוך () זוהי אינה נקודת עיניין.

# רגיסטרציה

רגיסטרציית תמונות, הינו תהליך להתאמת (Overlaying) תמונות שנלקחו מאותו הנוף, אך בזמנים שונים, נקודות מבט שונות או מצלמות שונות. תהליך הרגיסטרציה מיישר התמונות זו על זו על ידיי טרנספורמציה גיאומטרית כך שתמונה אחת נבחרת בתור תמונת הרפרנס, אליה נתאים את התמונה השניה, הנקראת גם תמונת המטרה.

ישנן שתי שיטות עיקריות לביצוע רגיסטרציה, לכל אחת מהן יתרונות וחסרונות, ובחירת השיטה המתאימה תלויה בהגדרת הבעיה לשמה הרגיסטרציה נחוצה:

**רגיסטרציה מבוססת עוצמה (Intensity based) -** שיטה זו ממפה כל פיקסל בתמונת המטרה לפיקסל המתאים בתמונת הרפרנס, לפי תבניות עוצמה בתמונה. שיטה זו הינה יותר כללית ומתאימה למגוון רחב של תמונות, בנוסף לא צריכה לקבל קלטים של נקודות עניין, ומתבצעת בצורה אוטומטית.

**רגיסטרציה מבוססת מאפיינים (Featured based) –** שיטה זו מקבלת מספר פיקסלים (features) בתמונה, מוצאת את הטרנספורמציה שלהם אל תמונת הרפרנס, ולאחר מכן מטילה על כל שאר הפיקסלים בתמונה את הטרנספורמציה שנמצאה. שיטה זו יותר מכוונת, ומוצאת את ההתמרה האופטימלית לנקודות עניין שבחרנו, וסבילה יותר לתבניות רקע דומות שעלולות לפגוע בטרנספורמציה מבוססת עוצמה.

**רגיסטרציה מבוססת תדר (PCR)** **-**  זו היא הרגיסטרציה אותה הטמענו בפרוייקט שלנו. PCR ((phase correlation registration, הינה שיטת עבודה בתחום התדר לשם שיערוך השהייה או תזוזה של שני העתקים השייכים לאותו האות, כפי שמתואר ב - [2].

טכניקה זו מבוססת על תכונות ההזזה של טרנספורם פורייה. ובאופן ספציפי יותר, נניח שני אותו דיסקרטים ומחזוריים עם התמרות פורייה  *בהתאמה.*

*מכאן שפונקציית הקרוס-ספקטרום המנורמלת של*  הינה-

נשים לב כי עבור כל תדר מתקיים :  *, בנוסף, פונקציית הקורלציה התלויה בפאזה בלבד ((POC , , הינה התמרת פורייה ההפוכה של .*

*כעת, נניח ש- הינה גרסה מוזזת של* , כלומר , כאשר אינו ידוע *ואותו אנו רוצים למצוא.*

*מתוך תכונות התמרת פורייה נקבל - .*

*במקרה זה, קל לראות כי , וכן , כאשר היא פונקציית הלם של דיראק. מכאן שאפשר לחלץ בקלות את ההזזה ,* , על ידיי מציאת המקסימום של.

את השיטה הזו ניתן להכליל עבור תמונות משניים וכן שלושה מימדים, כפי שנעשה ב- [3].

ישנן מספר סוגים של טרנספורמציות גיאומטריות אותן ניתן למדל (איור 4) בהתאם למודל התנועה אותו מניחים, נסביר על חלקן:

**הזזה (translation) –**

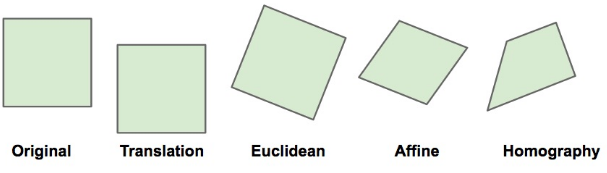
תמונת המטרה הינה הזזה של תמונת הרפרנס , כלומר יש רק שני פרמטרים לשערך -

**התמרה אפינית –**

התמרה זו כוללת הזזה, סיבוב ושינוי גודל (scaling) של התמונה, כלומר יש 6 פרמטרים לשערך (שלושה לכל ציר), התמרה זו שומרת על קווים מקבילים.

**התמרה הומוגרפית –**

התמרה זו כוללת את כל השינויים של ההתמרה האפינית, אך בנוסף יש לה גם מאפיינים מהעולם התלת-ממדי, כלומר אינה שומרת על קווים מקבילים. ריבוע העובר דרך התמרה זו, יכול להפוך לכל מרובע כלשהו. כלומר יש לשערך 8 פרמטרים בהתמרה זו (היא נכונה עד כדיי scaling)



איור 4: התמרות גיאומטריות שונות על ריבוע

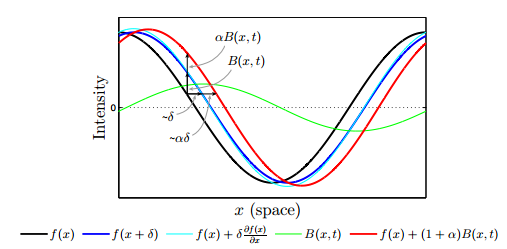
# החסרת תמונות בזמן

אחת הגישות לעקיבה אחר אובייקט דינאמי, או למציאת נקודות עניין עבור עקיבה הינה החסרת תמונות בזמן. ההנחה היא שבין פריים לפריים התמונה אינה משתנה (כמעט) ורק האובייקט שאנו מנסים לעקוב אחריו זז מעט –

כאשר השוויון השני הוא מפיתוח טור טיילור מסדר ראשון.

כעת, העברת אגפים לביצוע ההחסרה בין פריימים עוקבים תניב את התוצאה הבאה -

כלומר ההחסרה עבור פיקסלים בהם היה שינוי גדול במיקום (נגזרות המיקום גדולות) תיתן ערכים גבוהים עבור אותם הפיקסלים ובאופן זה נוכל למצוא את מיקום האובייקט.



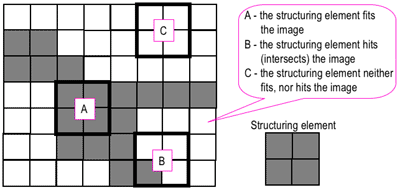
איור 5: ***המחשה לכך כי שינוי בגוון של פיקסל לאורך זמן יכול להעיד על תזוזה של פיקסל זה במקום*.**

באיור 5, ניתן לראות את אות הכניסה בזמן

ובזמן , . הינו הפרש הגוון בין הזמנים השונים, עבור הפרש גדול יותר נקבל תזוזה מרחבית גדולה יותר , עפ"י [4]

# פעולות מורפולוגיות על תמונות

פעולות מורפולגיות הן טכניקות לניתוח מבנים גיאומטריים. פעולות אלו מאוד נפוצות בתחום עיבוד התמונות, הן ממומשות על ידיי פונקציות לא ליניאריות ומופעלות על תמונות בינאריות. בפעולה מורופולוגית אנו דוגמים חלון קטן מהתמונה על ידי תבנית (structing element) ומחפשים קישוריות בין החלון לתבנית באופן הבא –



איור 6: שלושת סוגי התאמה של תבנית לתמונה

לאחר מכן מחליטים על מוצא הפונקציה לפי ההתאמה שהייתה בשלב הקודם.

פעולות מורפולוגיות בסיסיות

סימונים : A – Binary image in E , B – Structing element , E -Integer grid

* **Erosion**

כלומר התוצאה תיתן לנו '1' בכל הקואורדינטות (x,y) בהן התבנית (structing element), תואמת בדיוק את החלון אותו דוגמים, חלון A בדוגמה שמוצגת באיור 6 , ואפס אחרת.

פעולת ה erosion מכווצת את האלמנטים בתמונה על ידי כיווץ של הגבולות החיצוניים והפנימיים של כל צורה בתמונה

* **Dilation**

כלומר התוצאה תיתן לנו '1' כאשר יש חפיפה (מספיק אפילו פיקסל אחד) בין התמונה לתבנית , חלון B בדוגמה שבאיור 6.

פעולת ה dilation הינה הפעולה הדואלית לפעולת ה erosion, והיא בעצם מרחיבה את האובייקטים בתמונה על ידי הרחבת הגבולות הפנימיים והחיצוניים שלהם.

* **Opening**

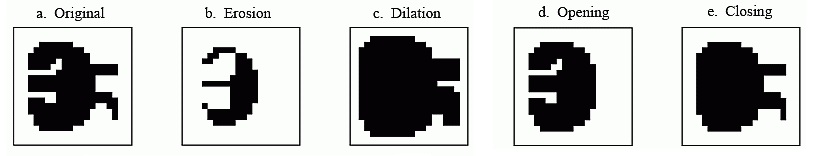
כלומר, פעולת ה opening היא שילוב של שתי הפעולות erosion ו- dilation כך שקודם מפעילים על התמונה erosion ולאחר מכן, על התוצאה מפעילים dilation.

פעולת ה opening מנתקת קשרים בין אובייקטים שהיו מחוברים בגשר דק של פיקסלים על ידי ה – erosion , ולאחר מכן, אובייקטים ששרדו את ה-erosion חוזרים לגודלם המקורי על ידיי פעולת ה dilation.

* **Closing**

פעולת ה closing היא שילוב של שתי הפעולות erosion ו- dilation כך שקודם מפעילים על התמונה erosion dilation ולאחר מכן, על התוצאה מפעילים dilation .

פעולת ה closingהינה הפעולה הדואלית לפעולת ה opening, והיא נקראת כך מכיוון שהיא סוגרת חורים באובייקטים ושומרת על הגודל המקורי.



איור 7: הפעלת פעולות מורפולוגיות בסיסיות

# Speeded up robust features (SURF)

Speeded up robust features (SURF) הינה שיטה לגילוי ותיאור נקודות עניין בתמונה, כפי שמתוארת ב - [5]. שיטה זו שימושית במגוון מטלות בתחום הראייה הממוחשבת כגון, זיהוי האובייקט ורגיסטרציה.

השיטה מתבססת על שיטת מציאת המתארים (descriptors), [scale-invariant feature transform](https://en.wikipedia.org/wiki/Scale-invariant_feature_transform) (SIFT) , אך עולה עלייה בביצועים ובמהירות החישוב.

לצורך מציאת נקודות עניין בשיטה SURF, נעשה שימוש בקירוב הדטרמיננטה של המטריצה האסיאנית:

כאשר הינה תוצאת הקונבולוציה בין מסכת הנגזרת הגאוסייאנית מסדר שני (LoG) עם התמונה בנקודה .

מציאת המתארים עבור נקודות העניין נעשית בשני שלבים, מציאת האוריינטציה של נקודת העניין במטרה להשיג עמידות של המתאר בפני שינויי סיבוב, ולאחר מכן חילוץ המאפיינים הסביבתיים באמצעות סכום התגובות של פונקציית Haar wavelet.

בכדי לתאר את הסביבה של נקודת העניין, מחולץ אזור ריבועי בגודל 20X20 פיקסלים ממורכז סביב נקודת העניין ומסובב בשיעור האוריינטציה אותו מצאנו בשלב הקודם. אזור ריבועי זה מחולק לתתי אזורים בגודל 4X4 כל אחד, עבור כל אחד מהם מחושבת ונשמרת התגובה לפונקציית Haar wavelet. בנוסך מתבצע משקול של התגובה הנ"ל ע"י פונקציה גאוסיאניית.

בכדי להתאים נקודות עניין בין תמונות שונות, נחשב את סכום ריבועי המרחקים (SSD) בין המתארים של נקודות המועמדות. ההתאמה הטובה ביותר מתקיימת בין הנקודות אשר המרחק בין המתארים שלהם מינימלי:

כאשר C קבוצת כל הנקודות המועמדות להתאמה, q נקודת העניין לה נרצה למצוא התאמה, D המרחק על פי המטריקה הנבחרת ו- הנקודה בעלת המרחק המינימאלי.

# רשתות נוירונים מלאכותיות

רשת נוירונים, כפי שמתוארת ב – [6] ,הינה מודל לעיבוד אינפורמציה. ההשראה למודל זה הינה הרשת העצבית הביולוגית הקיימת אצל בני אדם. כמו ברשת העצבית הביולוגית, כך גם רשת הנוירונים המלאכותית מורכבת מהרבה נוירונים (פרספטורים במקרה המלאכותי) המקושרים זה לזה.

מודל זה, בדומה למוח האנושי, לומד על ידי דוגמאות ומפיק מהן, על ידיי זיהוי מאפיינים, מה צריכה להיות התוצאה.

רשתות נוירונים מלאכותיות משמשות כיום לעיבוד אינפורמציה בהרבה מאוד תחומים בהם צריך לקבל החלטות על סמך דוגמאות קודמות:

* זיהוי של תבנית – אם זה מידע בכל צורה שהיא (טקסט, תמונה, נתונים מספריים)
* סיווג של תמונות
* זיהוי מגמות שונות במידע

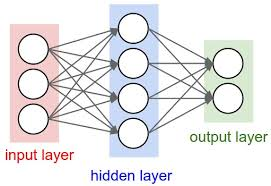
חלק מיתרונותיה של רשת הנוירונים המלאכותית על דרכי החישוב הקונבנציונליות שהיו בשימוש עד היום:

* למידה אדפטיבית - יכולת ללמוד איך לבצע משימות מסוימות תוך התבססות על מידע לאימון או ניסיון וללא אלגוריתם נתון.
* ארגון עצמי – רשת הנוירונים המלאכותית מארגנת את עצמה ואת דרך ההצגה וההתייחסות שלה למידע שהיא מקבלת במהלך שלב הלמידה.

רשת נוירונים ממוצעת מורכבת ממספר שכבות, לפחות 3, המאפשרות למידע לעבור מספר תהליכי עיבוד שונים לפני ההחלטה במוצא -

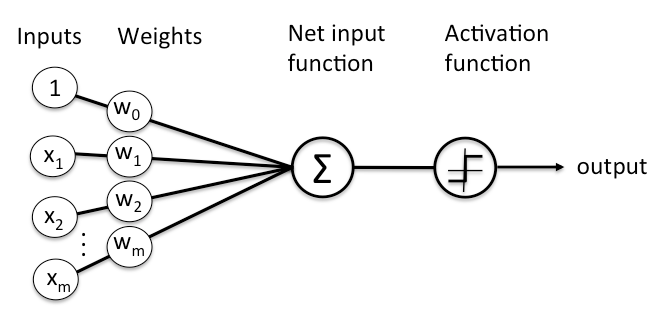
השכבה הראשונה היא שכבת כניסת הנתונים – כל נתון במידע שקיבלה הרשת נכנס לשכבה זו. השכבה האחרונה היא שכבת המוצא – שם בסופו של דבר מקבלים את ההחלטה של הרשת, בחלק מהמקרים מפעילים סף על שכבה זו על מנת לקבל תשובה בינארית של כן או לא.

בין השכבה הראשונה לאחרונה נמצאות השכבות החבויות, הן נקראות כך כיוון שהן לא חלק מממשק המשתמש, ושם נעשה בעצם רוב החישוב. ככל שיש יותר שכבות חבויות, כך הרשת יותר מורכבת והקישוריות בין המאפיינים של הכניסות יותר רחבה ומאפשרים למידה וביצועים טובים יותר בדרך כלל.



איור 8: דוגמה לרשת נוירונים בעלת שכבה חבויה אחת

כל שכבה ברשת בנויה ממספר נוירונים. לכל כניסה xבנוירון יש משקל w, אותה הרשת מכוונת תוך כדי הלמידה שלה כדי לדעת כמה חשובה אותה כניסה לנוירון הספציפי כדי לקבל את התפקוד הרצוי.



איור 9: המחשה לנוירון יחיד במערכת

המוצא של הנוירון ייקבע לפי סף אותו תקבע הרשת גם כן במהלך הלמידה באופן הבא –

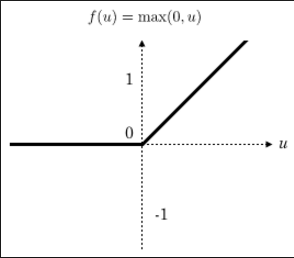
עדכון המשקולות נעשה על ידיי תהליך הנקרא Backpropogation , [7]. בתהליך זה מחושב המשקל של כל כניסה לנוירון שייתן שגיאה מינמיאלית במוצא בהתחשב במצב הרשת המאומנת עד כה. תהליך זה קורה רק בשלב הלמידה של המערכת.

# Alex-Net

AlexNet, כפי שהוצגה ב - [8], הינה רשת נוירונים מבוססת קונבולוציה (CNN) אשר אומנה על 1.2 מיליון תמונות ממאגר התמונות ImageNet LSVRC-2010 אשר מכיל 1000 מחלקות שונות. בתחרות ILSVRC-2012 זכה מודל זה במקום הראשון והשיג אחוזי שגיאה נמוכים מהרגיל, רק 15.3% שגיאה. AlexNet נחשב למודל פורץ הדרך לשימוש ברשתות נוירונים לפתרון בעיות מורכבות, כגון זיהוי תמונות.

מבנה נוירון (היחידה הבסיסית)

פונקצית האקטיבציה אשר מופעלת על המידע הנכנס לכל נוירון הינה nonlinearity as Rectified Linear Unit (ReLU):

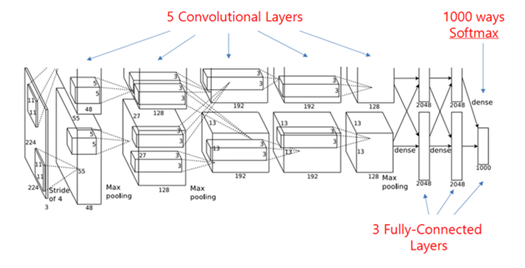


איור 10: פונקציית אקטיבציה מסוג RELU

השימוש בפונקציה זו מקטין משמעותית את זמן האימון של הרשת ובו בעת אינו פוגם בביצועי המודל.

מבנה הרשת

הרשת מורכבת משמונה שכבות לומדות, חמשת הראשונות הינן שכבות קונבולוציה ואילו שלושת האחרות שכבות fully-connected. מוצא שכבת ה- fully-connected האחרונה מוזנת כקלט לשכבת softmax בעלת 1000 נתיבי החלטה (מקבילים ל-1000 המחלקות השונות אותן אומנה הרשת לסווג). מוצאי שכבות הקונבולציה מוזנות לשכבות נורמליזציה וmax-pooling- כך ,שבסך הכל כוללת הרשת 25 שכבות המורכבות מ-650,000 נוירונים ובעלת מעל ל-60 מיליון פרמטרים שונים.



איור 11: מבנה הרשת, הכוללת חמש שכבות קונבולוציה, 3 שכבות fully-connected המלוות בשכבת softmax.

הרשת מקבלת כקלט תמונות בגודל 224X224 בשלושה ערוצי צבע (RGB). שכבת הקונבולוציה הראשונה מפעילה על הקלט 96 גרעינים שונים בגודל 11X11X3, כאשר צעד ההתקדמות של כל גרעין (stride) הינו 4 פיקסלים. לאחר נרמול ומעבר דרך שכבת pooling נכנס המידע לשכבת הקונבולוציה השנייה אשר מסננת את המידע על ידי גרעינים מגודל 5X5X48. השכבה השלישית, הרביעית והחמישית מחוברות אחת אל השנייה ישירות (ללא מעבר דרך נורמליציה וpooling-), כאשר גודל הגרעינים בכל שכבה הינו 3X3X256, 3X3X192 ו- 3X3X192 בהתאמה. כל אחת משכבות ה- fully-connected מכילה 4096 נוירונים.

# מודל SVM

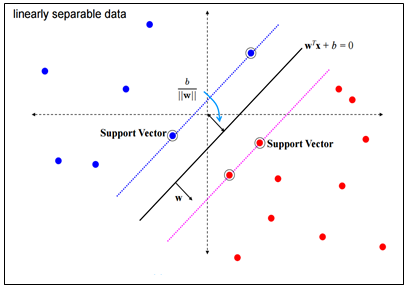
SVM או Support Vector Machine הוא אלגוריתם למידת מכונה Machine learning)) מבוקר המשמש לניתוח נתונים וכן משמש רבות למטרות סיווג, כפי שמתואר ב- [9]. למודל שני שלבים, שלב האימון שאותו נדרשים לבצע פעם אחת לקבלת המסווג שיפריד באופן הטוב ביותר בין סוגי המחלקות השונים, ושלב הסיווג, בו משתמשים כדי לסווג נקודה חדשה למחלקה מסוימת.

בעיית הלמידה של מודל ה-SVM מוגדרת באופן הבא:

קיים קשר לא לינארי שאינו ידוע , בין , וקטור הכניסה ממימד גבוה, ובין , סקאלר המוצא (או וקטור, במקרה של SVM עם מספר מחלקות). המידע היחיד שקיים על הקשר הוא מאגר הנתונים ללמידה , כאשר הוא מספר הנתונים במאגר. מוגדר להיות הערך הרצוי (הסיווג הנכון למחלקה), לכן SVM משתייך לטכניקות לימוד מבוקרות.

מודל זה מקבל וקטורי אינפורמציה המכילים מדידות של מאפיינים הנקראות features , כל feature כזה מקבל מימד משל עצמו לצורך ההפרדה בין אובייקטים עם מאפיינים שונים. וכגודל מספר ה features אנו בונים מרחב מאותו מימד.

כל אובייקט שהמודל מקבל כ Data ממוקם כנקודה במרחב ה-n מימדי בקואורדינטות המתאימות ל features ששייכים לו. לאחר הכנסת כל המידע למרחב, המודל מחפש את המישור, הנקרא hyper plane שמפריד את המרחב עם גבולות בין כל class אותו נרצה לסווג.



איור 12: דוגמה למציאת מישור חלוקה עם מימוש מודל SVM

בדוגמה הנראית לעיל , המרחב הוא דו ממדי, וקיימים שני features, וכן שני classes – באדום ובכחול.

ה- הינן הדגימות הקיצוניות ביותר, שיושבות על שוליי ההפרדה ותוחמות את אזורי ההחלטה.

במהלך שלב האימון, המודל מחשב פרמטרים , וכן היסט כדי ליצור את פונקציית ההחלטה שלו עבור כל כניסה –

על מנת ליצור את פונקציית ההחלטה האופטימלית, נפעל באופן הבא –

מכיוון שמשוואת המישור המפריד הינה , מתקיים גם ש -

לכן נוכל לבחור את הנרמול של כפי שנרצה. נבחר את הנרמול באופן הבא:

כאשר הם ה השלילי והחיובי בהתאמה.

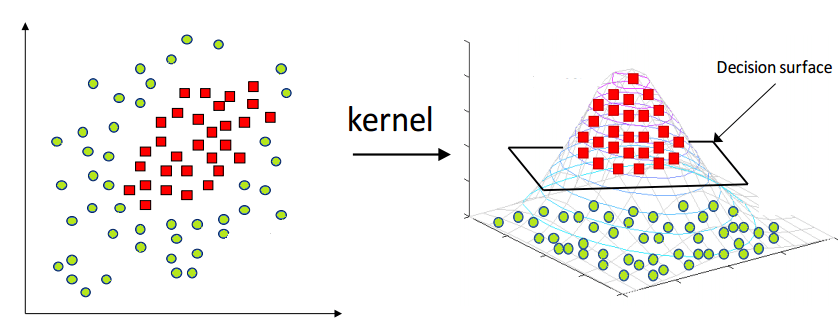
כעת, רוחב השוליים יחושב באופן הבא:

כלומר, על מנת להגדיל את רוחב השוליים (ובכך לקבל הפרדה טובה יותר) יש להקטין את הפרמטרים .

שלב אימון מודל ה –SVM יתבצע באופן הבא:

לאחר שלב הלמידה, במקרה הדו-מחלקתי ההחלטה תתבצע לפי פונקציית ההחלטה שהוגדרה לעיל באופן הבא –

מפריד ליניארי מניב תוצאות לא רעות, אך לפעמים קיימים מפרידים טובים יותר אך לא לינארים, על מנת למצוא אותם משתמשים ב Kernel Trick – מהלך זה מעתיק את דוגמאות האימון מהמרחב הלינארי המקורי למרחב ממימד גבוה יותר, מתוך הנחה שבמרחב החדש ימצא מפריד לינארי טוב יותר מאשר במרחב המקורי. ההעתקה נעשית בעזרת הטריק - מחליפים את המכפלה הפנימית בה משתמשים עבור המפריד הלינארי, בפונקציית kernel אשר מדמה את פיזורם מחדש של הווקטורים המקוריים במרחב עשיר יותר. כעת, המפריד הלינארי במרחב הוקטורי החדש הוא מפריד לא לינארי במרחב המקור.



איור 13: פעולת ה Kernel Trick

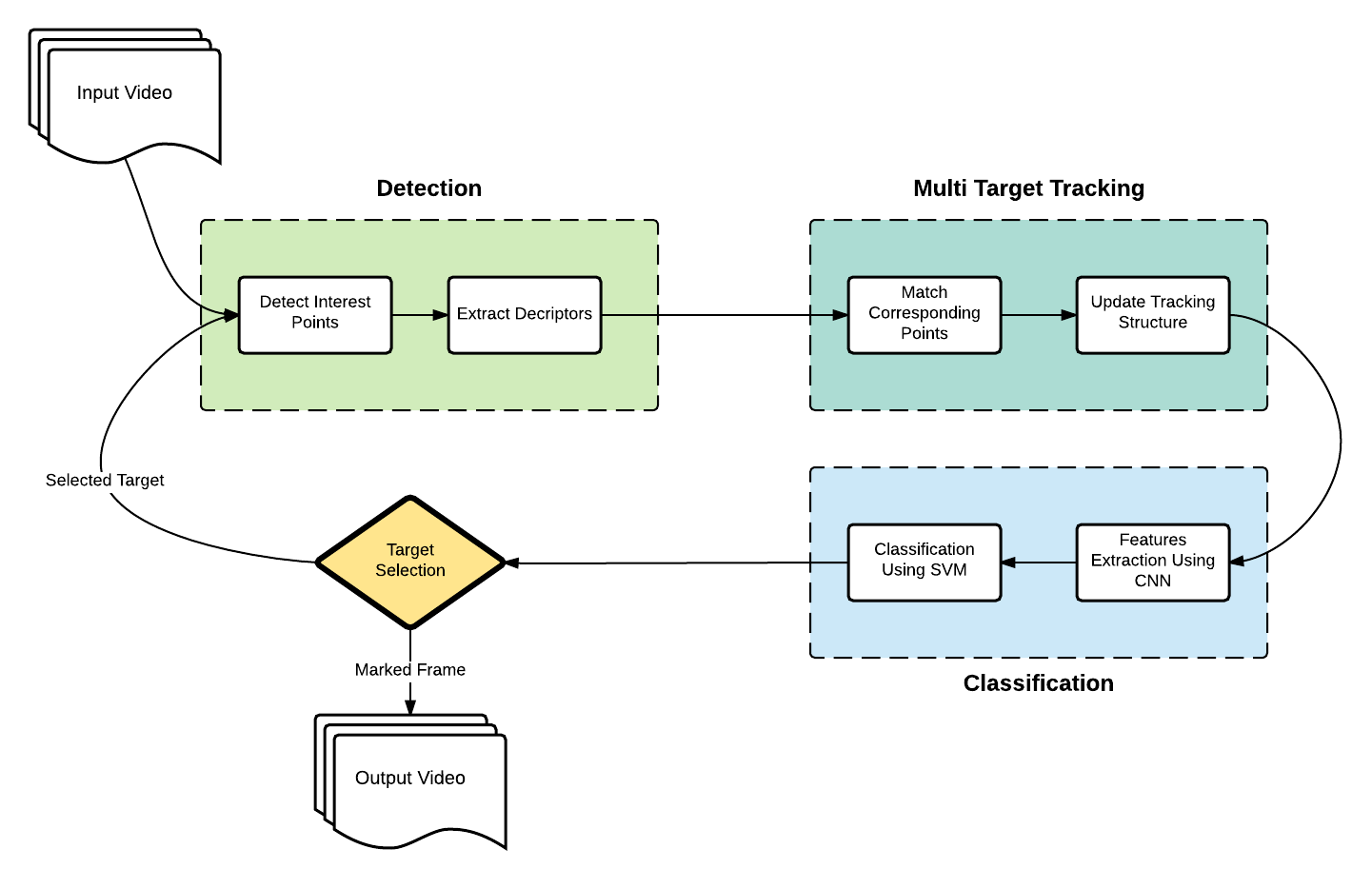
# **תכנון ומימוש**

כקלט מקבל האלגוריתם וידאו, ואזור כללי נתון בו ידוע לנו שקיים בו הרחפן והמטרה שלנו היא ראשית לגלות אותו , לבצע עליו עקיבה וכן בהמשך לאמת שמזוהה רחפן. אחד האתגרים הוא גילוי הרחפן, אשר לעיתים קטן מאוד וקשה לאתר אותו בתמונה. במיוחד על רקע נוף כגון עצים סבוכים, עננות, שדה חרוש וכו'. כמו כן, הרחפן גם משנה את צורתו, גודלו ומסלולו במהלך הצילום, כך שלא ניתן לחזות בוודאות על סמך נתונים רדיומטרים ופיזקלים היכן יהיה בפריים הבא.

העבודה מתבצעת על כל פריים בצורה סדרתית. תוך התחשבות במאפייני הרחפן, נחלץ תחילה נקודות עניין במספר שיטות המתבססות על עקרונות שונים במטרה לכלול את הרחפן בקבוצת נקודות אילו. בכך נצמצם את היקף הבעיה, מחיפוש בפריים כולו, למספר אזורים בולטים.

על מנת שנוכל להתאים נקודות אילו לאובייקטים הקיימים, נתאר אותן באמצעות סביבתן, ובכך נוכל לייצר עקיבה רציפה ולאורך זמן. כעת, כדי להבטיח כי האובייקט אחריו אנו עוקבים אכן רחפן.

# 4.1. דיאגרמת בלוקים



איור 14: דיאגרמת בלוקים מפורטת

בדיאגרמת הבלוקים באיור 14 לעיל, ניתן לראות את שלושת חלקי הפרויקט העיקריים –

* **Detection** - בחלק זה מקבלים מסגרת (פריים) חדשה עבורה מחלצים נקודות עניין לעקיבה וכן מאפיינים אותן באמצעות descriptors לשם חיפוש ההתאמה בשלב הבא.
* **Multi Target Tracking** – בחלק זה מקבלים את הנקודות ביחד עם ה- descriptorsשלהן, מתאימים בין נקודות מהפריים הנוכחי לנקודות מפריים קודם, וכן מעדכנים את מבנה הנתונים השומר את תהליך העקיבה עבור כל אובייקט.
* **Classification** – בחלק זה מקבלים את מבנה הנתונים המעודכן לפריים הנוכחי, ממיינים את המטרות ומסווגים אותם לפי רלוונטיות (רחפן או לא).

לאחר סיום שלושת השלבים, המטרה שלנו זוהתה כרחפן, מסמנים אותה בפריים הנוכחי עבור וידיאו המוצא וממשיכים באופן איטרטיבי לפריים הבא.

# 4.2. זיהוי ואפיון נקודות עניין

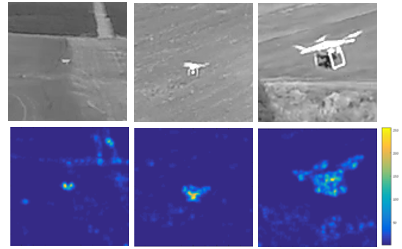
כיוון שאנו עובדים בגישת ה Sparse Optical Flow, יש צורך לאסוף נקודות עניין חשובות בתמונה, ביניהן הרחפן (או מספר נקודות מהרחפן) כדי לבצע את העקיבה.

על מנת למצוא את נקודות העניין השתמשנו בשתי שיטות אשר השלימו האחת את השנייה, וביחד הפיקו לנו את המידע הנחוץ על התמונה –

# 4.2.1. Harris Corner Detector

אלגוריתם זה מאפשר לנו למצוא נקודות בתמונה של אובייקטים בעליי טקסטורה מרחבית. כיוון שהרחפן הינו בעל טקסטורה שונה מזו של הרקע בו הוא נמצא, אלגוריתם זה יניב לנו נקודות עניין אשר נמצאות על הרחפן, ובאופן זה יאפשר לנו להמשיך ולעקוב.

לאחר חישוב תמונת Harris – כלומר תגובת כל אחד מהפיקסלים לאלגוריתם, מקבלים תמונה עם הרבה מידע אך רובו לא נחוץ. כיוון שאנו מחפשים רק את המקומות בהם יש מקסימום מקומי, שכן יש סיכוי רב יותר ששם נקודות הרחפן נמצאות, אנו מפעילים את אלגוריתם ה- non-maximum suppression , שפרט לנקודות המקסימום הלוקאליות, מאפס את כל שאר התגובות, כלומר כל הערכים המשניים שיכולים להפריע לחישוב. הנקודות שלא סוננו הופכות להיות נקודות העניין החדשות לתהליך עיבוד הפריים הנוכחי.



איור 15: בשורה העליונה דומאות לתמונות הרחפן, ובשורה השניה פלט אלגוריתם harris בהתאמה.

באיור 15 ניתן לראות שהאלגוריתם מתמודד היטב עם מקרים בהם גודל הרחפן שונה. הנקודות הצהובות הן בעלות התגובה החזקה ביותר, ואכן הן תואמות את מיקום הרחפן. דוגמאות אילו מוכיחות כי התבססות על טקסטורת הרחפן ביחס לרקע הינה מוצדקת.

# 4.2.2. חיסור תמונות בזמן

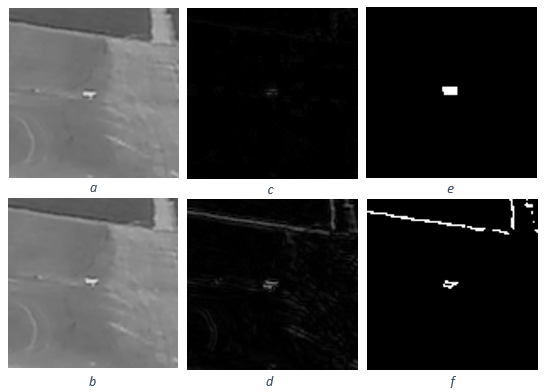
אלגוריתם זה מאפשר לנו למצוא אובייקטים נעים בתמונה. כיוון שהרחפן זז ביחס לרקע, הפרש של שתי תמונות עוקבות יניב נקודות עניין השייכות לרחפן.

על מנת שנוכל להחסיר בין שתי תמונות, ראשית יש צורך בביצוע רגיסטרציה, כלומר מציאת ההתמרה המתאימה, במקרה שלנו בחרנו להשתמש ב – PCR כדי למצוא את ההתאמה מהפריים הנוכחי אל הפריים הקודם. ביצוע רגיסטרציה מתבקש היות ובין פריים לפריים כיוון הצילום משתנה ולכן גם הרקע משנה את מיקומו. חיסור התמונות ללא הפעלת ההתמרה יניב תוצאה רועשת (איור d16).

לאחר הפעלת הרגיסטרציה, שתי התמונות מיושרות וכעת ניתן להחסיר ביניהן ולצפות שנקבל את מיקום הרחפן, שכן ההנחה היא שהוא האובייקט הנע הבולט ביותר (איור c16).

את תוצאת ההחסרה מעבירים דרך סף של גווני אפור, על מנת לקבל תמונה בינארית (איור e16). לעיתים, פרט לרחפן ישנן נקודות לא רצויות שעברו את הסף ומיוצגות בתמונה הבינארית, לכן אנו מפעילים מספר פעולות מורפולוגיות שמטרתן להעלים אובייקטים קטנים מהתמונה ולחזק אובייקטים בולטים.

לבסוף, מרכזי האובייקטים שנותרו בתמונה הבינארית הופכים להיות נקודות העניין החדשות שלנו ומתווספות לנקודות העניין שהגיעו מהשיטה הקודמת.



איור 16: הדגמת פעולת ההחסרה עם ובלי רגיסטרציה.

באיור לעיל, מוצגים שני פריימים עוקבים a ו-b. תמונה c הינה תוצאת החיסור לאחר רגיסטרציה ואילו תמונה d תוצאת החיסור ללא רגיסטרציה. תמונות e ו- f, הינן בינאריזציה של התמונות c ו-d בהתאמה. כפי שניתן לראות, ביצוע החיסור לאחר הרגיסטרציה הניב תמונה המכילה רק את הרחפן.

שילוב שתי השיטות מנצל את המידע הקודם שיש לנו על הרחפן. באמצעות ידע זה אנו מקבלים נקודות עניין חזקות הבלוטות הן בטקסטורה המרחבית והן בתנועתן על גבי הרקע, כלומר הן בעלות פוטנציאל גבוה להיות נקודות המטרה.

# חילוץ מתארים לאפיון

לאחר קבלת נקודות העניין, יש צורך לתאר אותן באמצעות מאפיינים ייחודים להן, על מנת לאפשר התאמה מיטבית בין נקודות בפרימיים עוקבים, המתחשבת בסביבת הנקודה.

לשם כך, אנו משתמשים בגישת האפיון של מתארים (descriptors), אשר מאפיינת כל נקודה באמצעות סביבתה (פיקסלים שכנים בחלון 11x11). השיטה באמצעותה בחרנו לחלץ את המתארים הינה **SURF** - speeded up robust features.

שיטה זו עמידה בפני שינויי סיבוב וגודל, תכונה הכרחית למתאר במקרה שלנו, היות והרחפן משנה את גודלו בצורה מהירה וכן, עקב תנועת הרחפן והמצלמה, מתקבלת תנועה סיבובית של הרחפן.

שימוש במתארים דלילים יותר, כגון מרחק אוקלידי מינימאלי בין הנקודות המעמדות להתאמה, אינו אמין מספיק מכיוון שאינו מגלם בתוכו את תכונות הסביבה של נקודת העניין. על כן, פעמים רבות, תתקבל התאמה שגויה בין אובייקטים קרובים בעלי מאפיינים שונים לחלוטין.

# Multi Target Tracking

אחת הסוגיות המרכזיות בבעיית העקיבה הינה קבלת ההחלטה אחרי איזה אובייקט, מבין המועמדים, תתבצע העקיבה. על מנת לבצע החלטה זו בצורה מושכלת, נדרש לאסוף מידע רב לאורך זמן על כל מועמד. ניתוח המידע יתרום לאפיונו של האובייקט ודירוגו כמועמד המוביל לעקיבה.

לשם כך נגדיר מבנה נתונים, אשר יקרא בהמשך טרק, שיאגד בתוכו את כלל המידע על מועמד, יכיל מאפיינים לגביו ופרמטרים שיעזרו לנהל את מבנה הנתונים.

השדות הבולטים במבנה הנתונים בו בחרנו להשתמש הינם:

* – מסלול הטרק, וקטור המכיל את מקומי הטרק החל מרגע יצירתו.
* – זמן (מספר הפריים) היווצרות הטרק.
* – מספר ההופעות של הטרק, מספר הפריימים בו נמצא שיוך לטרק.
* – מספר ההיעדרויות הרצופות של הטרק (פריימים ללא שיוך).
* – רדיוס החיפוש ((gating radius, מגדיר את שטח החיפוש. שיוך יכול להתבצע רק למועמדים המוכלים בשטח זה.
* – מדד לעוצמת נקודת העניין כפי שמתקבל מתגובת האריס או תוצאת ההחסרה (איור 15 ו-16 בהתאמה).
* – המתאר העדכני של הטרק (מתוך נקודת העניין האחרונה ששוייכה)**.**

איור 17: המחשת מבנה הנתונים.

# שיוך נקודות תואמות

כאשר בידינו המידע מהפריים הנוכחי – נקודות העניין והמתארים התואמים, נרצה להתאים בצורה מיטבית נקודות אלו לטרקים הקיימים.

תהליך ההתאמה של נקודה לטרק מסוים מתבצע בשני שלבים:

* סינון נקודות ראשוני – סינון כל הנקודות שאינן מוכלות בתחום החיפוש (GR) של הטרק. כיוון שמרחקן האוקלידי ממיקום הטרק גדול מדיי, קטן בצורה משמעותית הסיכוי שנקודות אלו תואמות. סינון זה מתבסס על הנחה מוקדמת כי האובייקטים לא זזים בצורה משמעויות בין פריימים עוקבים.
* התאמה בין המתארים – כיוון שבחרנו להשתמש במתארים מסוג SURF אחת הדרכים לחישוב מרחק במרחב זה הינה מטריקת שורש סכום ריבועי מרחקים (SSD). המתאר בעל המרחק הקטן ביותר הינו בעל ההסברות הגבוהה ביותר להתאמה, ולכן הנקודה לה שייך מתאר זה תשויך לטרק.

במידה לא נמצאה התאמה, לא נשייך לטרק זה נקודה חדשה ומבנה הנתונים יתעדכן בהתאם.

# עדכון מבנה הנתונים

לאחר התאמת נקודות העניין בין שני פריימים עוקבים, מתבצע עדכון של כל טרק בהתאם לתהליך אותו עבר:

* במידה ושויכה לו נקודת עניין מהפריים הנוכחי – הטרק מקבל את הקואורדינטות המרחביות של הנקודה, המתאר שלה וכן מתעדכן הסיגנל. בנוסף, זמן הנוכחות שלו, , וכן מספר ההיעדרויות, , מתאפס.
* במידה ולא התבצע שיוך – האובייקט שאחריו עוקב הטרק אבד, לכן מגדילים את רדיוס החיפוש של הטרק הספציפי, מעדכנים את פרמטר ההיעדרות, , וכן מעדכנים פרמטרים נוספים שיעזרו לאפיון בהמשך. כאשר מגיע לסף מסויים,, הדבר מעיד על כך שהטרק לא מצא נקודות תואמות למשך מספר פריימים גדול, ומתבצעת מחיקה של הטרק.

|  |
| --- |
|  |

אלגוריתם 1: פסאודו קוד המתאר את תהליך עדכון מבנה הנתונים.

בסיום שלב זה, ממוינים הטרקים על פי משך קיומם, , כך שהטרק שקיים זמן רב יותר מקבל עדיפות גבוהה. מדד נוסף למיון הינו עוצמת הטרק, , כאשר גם במדד זה עוצמה גבוהה שוות ערך לעדיפות גבוהה יותר.

# סיווג מטרות

כעת, כאשר בידינו הטרקים המועמדים להיות נקודת המטרה, נרצה לוודא, בצורה מושכלת, כי הטרק המוביל הוא אכן רחפן. לצורך כך, בחרנו להשתמש ברשת נוירונים לצורך אפיון האובייקט ובמודל SVM לצורך סיווגו כרחפן.

כשלב מקדים אנו נדרשים לאמן את המודלים בהם נרצה להשתמש. לצורך כך נדרש מאגר נתונים גדול ומתויג. על מנת ליצור מאגר זה השתמשנו במספר סרטוני למידה מהם, בעזרת אלגוריתם העקיבה הראשוני (ללא השימוש במודלים לומדים), חילצנו תמונות של הרחפן והרקע הסובב אותו. תמונות אילו שויכו למחלקות שונות, רחפן ורקע, והוזנו למודל במהלך שלב האימון.

מאגר התמונות הכולל כ- אלף תמונות (כל מחלקה מכילה כ- אלף תמונות) בגדלים הנעים בין ל- פיקסלים, בפורמט RGB.

|  |
| --- |
| *C:\Users\rafael bouzaglo\Desktop\figures\dorneDB.PNGC:\Users\rafael bouzaglo\Desktop\figures\backgroundDB.PNG*  איור 18: דגימת מאגר התמונות ממחלקת הרחפן.  איור 19: דגימת מאגר התמונות ממחלקת הרקע. |

# חילוץ מאפיינים באמצעות CNN

חילוץ המידע החיוני לאפיון הרחפן מתמונה הינה מטלה מסובכת לאלגוריתם חישוב קונבנציונלי. כיום, שיטה נפוצה לפתרון בעיות דומות הינה שימוש באלגוריתמי למידה עמוקה, ופרט שימוש CNN.

בנייה ואימון רשת הינו תהליך הדורש זמן רב ומאגר נתונים בהיקף של מיליוני תמונות מתויגות. על כן, בחרנו להשתמש ברשת קיימת,Alex net, אשר אומנה לסווג כאלף אובייקטים (אך לא רחפנים). החלק הארי במבנה הרשת הינו חילוץ המאפיינים מהתמונות כך שכל מחלקה תתואר באופן נבדל משאר המחלקות. מפני שהרשת אומנה על מגוון רחב של תמונות טבעיות, יש לה את היכולת לחלץ מאפיינים טובים מתמונות בעלות אופי דומה.

המאפיינים אשר חילצנו הינם מוצאי שכבת ה-fc7 אשר הינה אחת השכבות האחרונות בתהליך חילוץ המאפיינים ברשת ולאחריה מתחיל תהליך הסיווג.

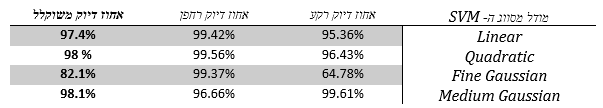
השימוש ברשת נעשה בשני שלבים בפרויקט:

* לשם אימון מודל ה-SVM (אודותיו נפרט בהמשך) יש צורך להמיר את מאגר התמונות שיצרנו למאגר של וקטורי מאפיינים אשר יאפשר למודל ה-SVM להפריד בצורה טובה בין מחלקות הסיווג השונות.
* כחלק מתהליך סיווג הטרקים, ניצור לכל טרק וקטור מאפיינים אשר מאוחר יותר יכנס אל מודל ה-SVM המאומן לצורך סיווגו כרחפן או כרקע.

# 4.4.2. סיווג באמצעות מודל SVM

לאחר חילוץ המאפיינים ע"י רשת הנוירונים נותר לקבוע האם הטרק המוביל הינו רחפן. לצורך כך נעשה שימוש במודל SVM.

תחילה נדרש לאמן את המודל על המידע הרלוונטי לנו. תהליך האימון נעשה ע"י שימוש ב-app – Classification learner של מטלב. לצורך השימוש בApp- זה, נדרש להכין את מאגר הנתונים (בצורת טבלה של מטלב) אשר כל שורה כוללת וקטור מאפיינים ואת התיוג למחלקה. לבסוף נבחר מודל הסיווג – Medium Gaussian SVM אשר הניב את ההפרדה הטובה ביותר בין מאפייני הרחפן לבין מאפייני הרקע.



טבלה 1: השוואת מסווגי SVM שונים.

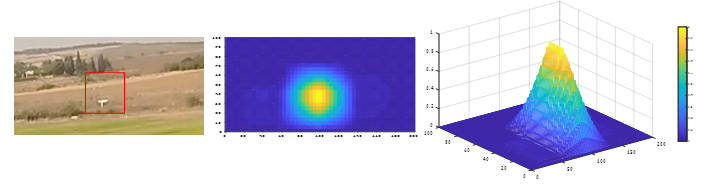
כפי שניתן לראות למודל הנבחר, Medium Gaussian, אחוזי הדיוק המשוקללים הגבוהים ביותר. זאת ועוד הוא בעל אחוזי דיוק בסיווג הרקע הגבוהים ביותר, תכונה חשובה לנו היות וזיהוי רקע כרחפן עלול לפגוע בהמשך תהליך העקיבה (בעוד זיהוי רחפן כרקע אמור להיות מגובה בתהליך העקיבה הראשוני).



טבלה 2: Confusion matrix של המודל הנבחר.

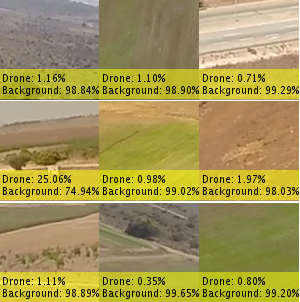
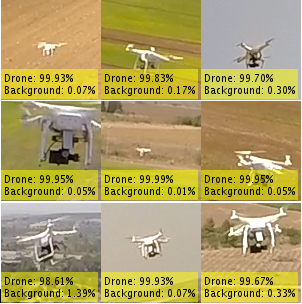
כעת, לאחר יצירת המודל המאומן, נשתמש בו לסיווג הטרקים המובילים. נזכור כי הטרקים המגיעים לשלב זה ממוינים על פי משך הזמן שהם קיימים ועל פי העוצמה שלהם בתמונות חילוץ נקודות העניין, על כן הנקודה הראשונה נחשבת למועמדת המובילה להיות נקודת המטרה, אחריה הנקודה השנייה וכן הלאה.

נרצה לוודא שהטרק המוביל אכן רחפן, לצורך כך נבדוק האם הסיווג של טרק זה הינו רחפן, במידה וכן נמשיך בשלבי האלגוריתם הבאים. אחרת, נבדוק את סיווג הטרק הבא אחריו. נבצע בדיקות אילו עד אשר נגיע לטרק אשר סווג כרחפן (לכל היותר 5 בדיקות) במקרה זה נעלה את הטרק שסווג כרחפן להיות המוביל ונמשיך בשלבי האלגוריתם הבאים. במידה ולא נמצא טרק שסווג כרחפן, נמשיך עם המיון הראשוני, והטרק בעל העדיפות הגבוהה ביותר יבחר כמטרה.



איור 20: תגובת המסווג לחלקים שונים בתמונה.

באיור 20 ניתן לראות את תגובת המסווג לתמונת הכניסה (התמונה השמאלית).התמונות האמצעית והימנית מייצגות תמונת התוצאה, במבט על וכמפת גבהים, בהתאמה. כל פיקסל בתמונת התוצאה הינה תגובת המסווג לחלון בגודל אשר מרכזו באותו הפיקסל (מעבר בשיטת sliding window). הגוונים הצהובים מעידים על הסתברות גבוהה לזיהוי רחפן באותו החלון. באיור השמאלי ניתן לראות את החלון אשר ממורכז סביב הפיקסל בעל התגובה החזקה ביותר (מסומן באדום).



איור 22: מדגם של תמונות רחפן עם תיוג המסווג

איור 21: מדגם של תמונות רקע עם תיוג המסווג

באיורים 21,22 ניתן לראות את אחוזי הדיוק הגבוהים של המסווג בהתאם לאובייקט בתמונה.

# ניתוח סיכום והצעות להמשך

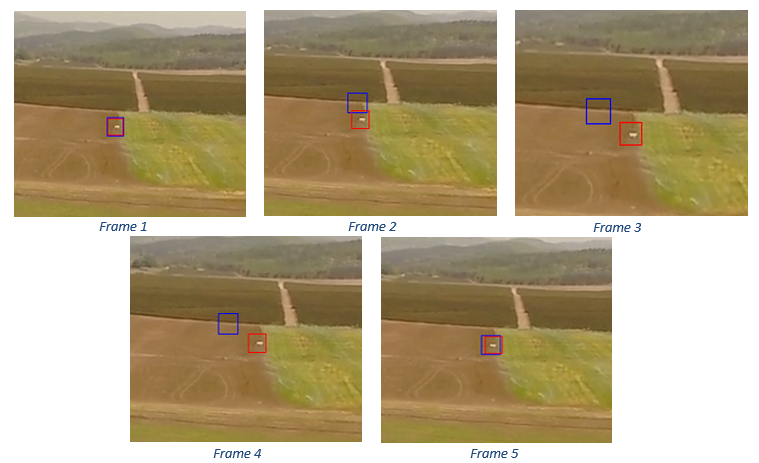
# 5.1. חשיבות שלב הסיווג לתהליך העקיבה

לצורך ניתוח תוצאות האלגוריתם ובדיקת התרומה של חלקו השני, הסיווג, ביצענו מספר בדיקות על האלגוריתם עם ובלי המסווג. השימוש באלגוריתם ללא המסווג הניב תוצאות עקיבה סבירות אך, לעיתים, לא רציפות.

תוצאות אילו מתקבלות מכיוון שמיון הנקודות נעשה על פי מדדי זמן ועוצמה בלבד (כפי שפורט בחלק 2.4.3 ) ועל כן, הרחפן אינו ממוין כטרק בעל העדיפות העליונה, ולא נבחר כמטרה. כתוצאה מכך מסרטון התוצאה נראה כי המטרה מרצדת בין אובייקטים שונים והעקיבה אחר הרחפן אינה רציפה.

במקרים מסוימים, עקב המצאות הרחפן ברקע בעל טקסטורה מורכבת, מסגל לעצמו האלגוריתם מטרה חדשה והעקיבה אחר הרחפן תמה.

עם שילובו של המסווג, מצאנו כי האלגוריתם בעל יכולות עקיבה טובות יותר, והינו עמיד בפני סיטואציות מורכבות. במקרים שבדקנו, רוב ככולם, התקבלו תוצאות טובות יותר מאילו שהתקבלו תוך שימוש באלגוריתם ללא מסווג.



איור 23: עקיבה עם מסווג (באדום) ובלי המסווג (בכחול).

כפי שניתן לראות באיור 23, בשני המקרים התקבלה עקיבה אחרי הרחפן, אך בעוד שהעקיבה עם המסווג הייתה יציבה לאורך כל הפריימים, העקיבה ללא המסווג לא הייתה רציפה והמטרה המובילה השתנתה מספר פעמים במהלך פריימים אילו.



איור 24: מיון הטרקים עם ובלי השימוש במסווג.

איור 24 ממחיש את יתרונות השימוש במסווג לצורך מיון הטרקים המובילים. בשורה העליונה ניתן לראות את הטרקים כפי שהתקבלו לאחר המיון הראשוני (לפני פעולת המסווג). ניתן לראות כי לפי מיון זה, מסדר טרקים לא נכונים (המכילים רקע בלבד) כעדיפים יותר ואילו הרחפן מוין כטרק הרביעי בעדיפותו. תחת כל תמונה בשורה זו ניתן לראות את תוצאת המסווג, כאשר ניתן לראות כי שלושת הטרקים הראשונים סווגו באופן נחרץ כרקע, ואילו הטרק הרביעי, הרחפן, סווג בוודאות גדולה כרחפן. בשורה התחתונה ניתן לראות את המיון החדש של הטרקים (לאחר הסיווג) ובו הרחפן דורג ראשון ואחריו נמצאים שאר הטרקים לפי סדר מיונם הקודם.

# 5.2. ניתוח תוצאות

|  |  |  |
| --- | --- | --- |
| **מספר סרטון הבוחן** | **תיאור הסרטון** | **אחוזי דיוק בעקיבה -** |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

# 5.3. סיכום

בפרויקט זה נחשפנו מורכבותה של בעיית עקיבה מאין זו, כאשר האובייקט והרקע משתנים במהירות ובסקאלה רחבה. תוך העמקה בבעיה זו, מצאנו לנכון לשלב מספר שיטות כאשר לכל אחת יתרונות והתייחסות למאפייני הבעיה בצורה שונה, כל שמכלול השיטות הניב תוצאה אופטימלית.

היעדים המרכזיים אותם הצבנו, תחילה עקביה סבירה אחר הרחפן תוך שימוש בשיטות עבודה המתבססות על אלגוריתמים מתחום הראייה הממוחשבת. לאחר השגת מטרה זו, שילבנו שיטות חדשניות, מימשנו מערכת לומדת ויצרנו מסווג ייחודי לבעיה שלנו. באמצעות מערכות אילו השגנו עקיבה רציפה וטובה אחר הרחפן, המטרה אליה שאפנו להגיע.

# 5.4. הצעות להמשך

באלגוריתם המוצע ניתן להטמיע שיפורים נוספים שיביאו לעליית ביצועים:

* בחינת שיטות נוספות למציאת נקודות עניין וחילוץ מתארים. במהלך עבודתנו בחנו את הוספת אלגוריתם fast-9 כתוספת למערכת מציאת נקודות העניין, אך לא ראינו שיפור בביצועים ולכן החלטנו שלא לצרפו. יתכן ואלגוריתמים המתבססים על שיטות שונות יניבו תוצאות טובות יותר.
* יצירת מסווג מורכב יותר, עם מספר גדול יותר של מחלקות זיהוי ( רכבים, ציפורים, כלי תעופה נוספים). מסווג זה יוכל לאפשר התמודדות עם מקרים מסובכים יותר לזיהוי, כאשר ישנם אובייקטים שדומים יותר לרחפן מאשר לרקע, ועל כן עלולים להיות מסווגים כרחפן.
* בחינת פתרון end-to-end תוך שימוש באלגוריתמי למידה עמוקה. ניתן לראות כי השימוש בפתרון מסוג זה נפוץ לבעיות דומות לבעיה איתה התמודדנו.

# **ביבליוגרפיה**

|  |  |
| --- | --- |
| [1] | C. Harris ו M. Stephens, “A Combined Corner And Edge Detector,” 1988. |
| [2] | A. Alba, R. M. Aguilar-Ponce, J. F. Viguera-Gomez and E. Arce-Santana, "Phase Correlation Based Image Alignment," *MICAI 2012: Advances in Artificial Intelligence,* pp. 171-182, 2012. |
| [3] | E. D. Castro ו C. Morandi, “Registration of Translated and Rotated Images Using,” *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE,* PAMI-9, NO.5, 1987. |
| [4] | H.-Y. Wu, M. Rubinstein, S. Eugene, J. Guttag, F. Durand and W. T. Freeman, "Eulerian Video Magnification for Revealing Subtle Changes in the World," *ACM Transaction on Graphics,* vol. 31, 2012. |
| [5] | H. Bay, T. Tuytelaars ו L. V. Gool, “SURF: Speeded Up Robust Features,” *ECCV 2006,* 2006. |
| [6] | L. Francis, “The Basic of Neural Networks Demystified,” *Contingencies,* November/December, pp. 56-61, 2001. |
| [7] | D. Rumelhart, G. Hinton ו R. J. Wiliiams, “Learning Internal Representaion by Error Propagation,” Chapter 8, pp. 317-362, 1985. |
| [8] | A. Krizhevsky, I. Sutskever ו G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” *Advances in neural information processing systems,* 2012. |
| [9] | D. Fradkin ו I. Muchnik, “Support Vector Machines for Classification,” *DIMACS Series in Discrete Mathematics and Theoretical Computer Science,* 2006. |

# **רשימת איורים וטבלאות**

[איור 1 : דיאגרמת בלוקים כללית 3](#_Toc485629550)

[איור 2: זיהוי רחפן בתנאי רקע קשים 4](#_Toc485629551)

[איור 3: גרף המתאר את איפיון הפיקסל אל מול הערכים העצמיים(harris) 6](file:///C:\Users\Student\Downloads\ספר%20פרויקט.docx#_Toc485629552)

[איור 4: התמרות גיאומטריות שונות על ריבוע 8](file:///C:\Users\Student\Downloads\ספר%20פרויקט.docx#_Toc485629553)

[איור 5: גרף של שינוי בגוון אל מול תזוזה במרחב 9](#_Toc485629554)

[איור 6: שלושת סוגי התאמה של תבנית לתמונה 9](#_Toc485629555)

[איור 7: הפעלת פעולות מורפולוגיות בסיסיות 10](#_Toc485629556)

[איור 8: דוגמה לרשת נוירונים בעלת שכבה חבויה אחת 12](file:///C:\Users\Student\Downloads\ספר%20פרויקט.docx#_Toc485629557)

[איור 9: המחשה לנוירון יחיד במערכת 13](#_Toc485629558)

[איור 10: פונקציית אקטיבציה מסוג RELU 14](#_Toc485629559)

[איור 11: מבנה רשת Alex-Net 14](#_Toc485629560)

[איור 12: דוגמה למספר מישורי חלוקה עם מימוש מודל SVM 15](#_Toc485629561)

[איור 13: פעולת ה Kernel Trick 16](#_Toc485629562)

[איור 14: דיאגרמת בלוקים מפורטת 17](#_Toc485629563)

[איור 15: תגובת אלגוריתם Harris לתמונות שונות. 18](#_Toc485629564)

[איור 16: הדגמת פעולת ההחסרה עם ובלי רגיסטרציה. 19](#_Toc485629565)

[איור 17: המחשת מבנה הנתונים. 21](#_Toc485629566)

[איור 18: דגימת מאגר התמונות ממחלקת הרחפן. 23](file:///C:\Users\Student\Downloads\ספר%20פרויקט.docx#_Toc485629567)

[איור 19: דגימת מאגר התמונות ממחלקת הרקע. 23](file:///C:\Users\Student\Downloads\ספר%20פרויקט.docx#_Toc485629568)

[איור 20: תגובת המסווג לחלקים שונים בתמונה. 25](file:///C:\Users\Student\Downloads\ספר%20פרויקט.docx#_Toc485629569)

[איור 21: מדגם של תמונות רקע עם תיוג המסווג 25](file:///C:\Users\Student\Downloads\ספר%20פרויקט.docx#_Toc485629570)

[איור 22: מדגם של תמונות רחפן עם תיוג המסווג 25](file:///C:\Users\Student\Downloads\ספר%20פרויקט.docx#_Toc485629571)

[איור 23: עקיבה עם מסווג (באדום) ובלי המסווג (בכחול). 26](#_Toc485629572)

[איור 24: מיון הטרקים עם ובלי השימוש במסווג. 27](#_Toc485629573)

[טבלה 1: השוואת מסווגי SVM שונים. 24](#_Toc485629932)

[טבלה 2: Confusion matrix של המודל הנבחר. 24](#_Toc485629933)