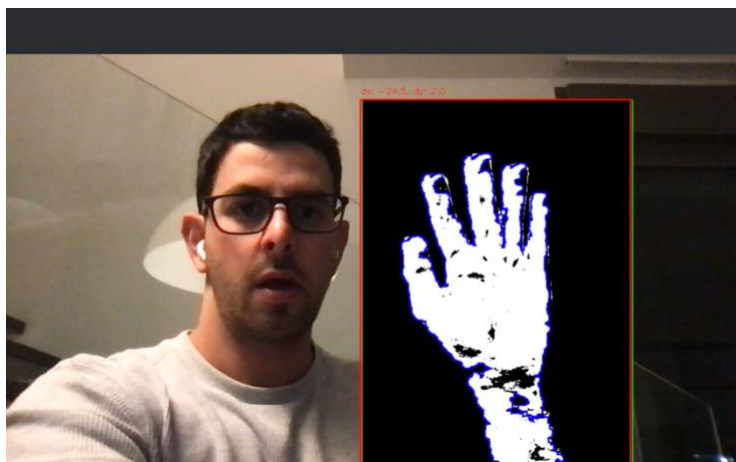


מבוא לעיבוד ספרתי של תמונה

דו"ח גמר

רחפן המוזז על ידי תנועות ידיים



מגישים:

רועי שחמון - ת"ז 206564759

עמית חי בוסאני - ת"ז 208384123

רואי גרייף - ת"ז 315111401

עמית אהרוני - ת"ז 206790149

תקציר הפרויקט:

אנו עסקנו בהמרת תנועות יד אנושית לפקודות המניעות רחפן בזמן אמת, תוך התבססות על כלים מעולמות הראייה הממוחשבת והסגמנטציה. בעזרת איתור מדויק של תווי היד וסינון צבע עור, היא מזהה ומנתחת מחוות שונות גם בתנאי תאורה ורעש משתנים. חיבור המערכת לרחפן מתבצע באופן ח' באמצעות ממשק תוכנה ייעודי, ומצלמת המחשב האישי משמשת כאמצעי התצפית המרכזי. כך מתקבלת דרך שליטה אינטואיטיבית וגמישה, ללא צורך בשלט או באביזרים ייעודיים.

הקדמה:

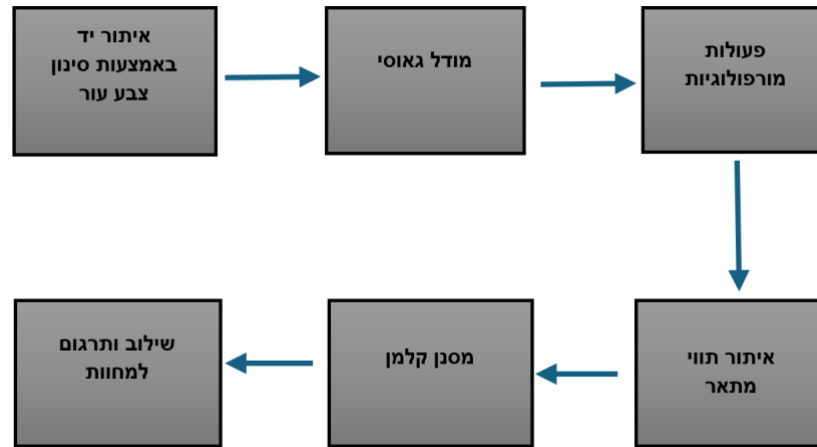
בפרויקט זה אנו מתמקדים בפיתוח מערכת לזיהוי, לעיבוד ולעקיבה אחר יד אנושית בזמן אמת, תוך שימוש בכלי עיבוד תמונה וראייה ממוחשבת.

המטרה המרכזית היא לאפשר זיהוי מדויק של היד בסביבתה הטבעית, להפרידה מהרקע, ואז לעקוב אחר תנועותיה באופן שוטף, כאשר השליטה והתנועה של הרחפן נשענות על מחוות היד וניתוחן. כך ניתן להפוך את היד לממשק משתמש טבעי ונטול מגע, המאפשר להטיס את הרחפן ולהוביל אותו לכיוונים שונים באמצעות תנועות יד פשוטות.

התוצאות המצופות כוללות זיהוי עקבי של היד גם בתנאי רקע מגוונים, ניתוח מדויק של תנועותיה וחיבור מוצלח של פקודות המחוות ישירות לרחפן בזמן אמת. האתגרים העיקריים טמונים בסגמנטציה מדויקת של היד, למידה ושיערוך של מיקומה ותנועתה, והתמודדות עם רעשים, תאורה משתנה ותנועות לא צפויות. אתגר נוסף היה להגביל את תנועת הרחפן שיתאים לספריה, לאזור מאוכלס יחסית. כמו כן, יש צורך בממשק תרגום יעיל הממיר את המחוות לתנועות הרחפן בפועל, ללא השהיות משמעותיות.

הטכנולוגיות בהן נעשה שימוש כוללות מעבר למרחב צבעים HSV, מסנני צבע עור, איתור תווי מתאר (Contours) ויישום מסנן קלמן (Kalman Filter) לייצוב התנועה ולעקיבה חלקה. המצלמה המשמשת לזיהוי היד היא המצלמה האישית המותקנת במחשב. בין האילוצים העיקריים של הפרויקט נכלל הצורך ברקע כללי שאינו משתנה באופן תדיר, וכן האפשרות להשתמש בשתי ידיים כרצוננו במהלך השליטה על הרחפן. שילוב כל הרכיבים הללו מאפשר לפתח מערכת אינטואיטיבית ואמינה, ההופכת את היד לממשק משתמש טבעי ויעיל לניווט הרחפן.

תיאור האלגוריתם:



איתור יד באמצעות סינון צבע עור - נלמד בהרצאות הראשונות על מרחבי צבע: באלגוריתם קיים טווח מוגדר לערכי גוון (Hue), רוויה (Saturation) ובהירות (Value) המייצגים עור אנושי במרחב HSV באמצעות הגדרת סף מתאים, נוצרת מסכה בינארית המפרידה את היד מהרקע, זהו השלב הראשון בזיהוי כף היד של המפעיל.

מודל גאומטרי: **(Mahalanobis Distance)** נלמד בנושא מספר 3 בהרצאות. אוסף פיקסלים של עור כף היד נאמד כהתפלגות נורמלית, כאשר הפרמטרים כמו ממוצע ומטריצת קו-וריאנצה, לאחר מכן, בכל פריים חדש מהמצלמה, מחשבים לכל פיקסל את מרחק מהלנוביס - זהו מרחק המודד עד כמה פיקסל מסוים קרוב לצבע העור כפי שהוגדר בדגימה, בהתחשב גם במידת הפיזור והצורה של ההתפלגות. אם המרחק נמוך מסף מסוים, סימן שהפיקסל דומה מספיק לצבע העור ונחשב חלק מאזור היד, אם לא, הוא נדחה כרקע.

חישוב וקטור הממוצעים - $\frac{1}{N} \sum_{i=1}^N X_i$; חישוב הממוצע של כל ערוץ צבע.

חישוב מטריצת הקו-וריאנצה, $\Sigma = \sum_{i=1}^N (X_i - \mu)(X_i - \mu)^T$

וקטור הפרש הינו $\Delta X = X_i - \mu$: מטריצה זו היא בגודל 3x3 כאשר האלמנטים על האלכסון הראשי מייצגים את השונות של כל ערוץ צבע, והאלמנטים מחוץ לאלכסון מייצגים את הקו-וריאנצה בין זוגות ערוצים.

חישוב מרחק מהלנוביס, $Dm(x) = \sqrt{\sum (\Delta X)^T (\Delta X)^{-1} (\Delta X)}$; כאשר ΔX הוא הטרנספוז של וקטור הפרש, ΔX^{-1} היא מטריצת הקו-וריאנצה ההפוכה.

סיכום: חישוב מרחק מהלנוביס עבור כל פיקסל מאפשר לקבוע עד כמה צבע הפיקסל דומה לצבע העור המודל, וכך אנו מעריכים את ההסתברות שכל פיקסל בתמונה שייך לאזור העור, תוך התחשבות בשונות ובקו-וריאציה בין ערוצי הצבע. שיטה זו מסייעת בזיהוי מדויק של אזורי עור בתמונה, גם בתנאים של תאורה משתנה ורעשים.

פעולות מורפולוגיות: פעולות מורפולוגיות (Erosion) ו: (Dilation)-

לאחר יצירת מסכה בינארית ראשונית) באמצעות סינון בצבע או חישוב מהלנוביס, (אנו מקבלים תמונה שבה יש אזורים לבנים המייצגים את היד, אך עלולים להופיע גם "כתמים" (או רעשים נקודתיים) למשל פיקסלים בודדים שזוהו בטעות כעור. (כדי לשפר את איכות המסכה, מופעלות שתי פעולות עיקריות:

1. **Erosion** כרסום): (אוכלת "מהאזורים הלבנים על-ידי הסרת שכבת פיקסלים דקה מהשוליים. כך, אם ישנם רעשים קטנים מאוד, הם ייעלמו) כי לעיתים קרובות הם אינם מספיק גדולים כדי לשרוד את פעולת ה. (Erosion-מנגד, אזור היד עצמו עשוי להצטמצם מעט, אך בדרכ כלל המסה המרכזית שלו נשמרת.

2. **Dilation** הרחבה): (מרחיבה את האזורים הלבנים וממלאת חורים קטנים שעלולים להופיע ביד) למשל אזור כהה בכף היד שלא זוהה כעור. (פעולה זו" משיבה "חלק מהפיקסלים שהוסרו בשלב ה Erosion- ומחברת אזורים סמוכים שהיו עלולים להתנתק. הנוסחה נראת מהצורה הבאה :

החזרה (Iterations) על Erosion/Dilation מספר פעמים נועדה לאזן בין הסרת רעשים לבין שמירה על צורת היד. הגודל והצורה של ה- Kernel למשל מטריצה 3×3 או 5×5 מכתבים עד כמה הפעולה תהיה "חזקה" ותשפיע על השוליים. התוצאה הסופית היא מסכה חלקה ורציפה יותר, שקל לאתר בה את תווי המתאר של היד.

איתור תווי מתאר (**Contours**) והפקת תכונות מהמתאר נלמד כנושא מספר 10 בהרצאות : לאחר שיצרנו מסכה בינארית" נקייה" "בזכות הסינון לפי צבע עור ופעולות מורפולוגיות, (אנו ניגשים לאתר תווי מתאר (Contours) בתמונה. הדבר מבוצע באמצעות הפונקציה `cv2.findContours` שמחזירה רשימה של אוספי נקודות שכל אחד מהם מייצג מתאר רציף של אזור לבן במסכה.

1. בחירת המתאר הגדול ביותר: במציאות, ייתכן יותר ממתאר אחד) למשל, רעשים קטנים שנתרו. (עם זאת, מניחים שהמתאר הגדול ביותר שייך לכף היד. האלגוריתם ממין את כלל המתארים לפי שטח (Area) ובוחר את זה בעל השטח המרבי.

-
2. חישוב תיבת החיבוק (**Bounding Box**) על המתאר הגדול ניתן להפעיל פונקציית `cv2.boundingRect` המחזירה ארבעה ערכים `x, y, w, h` – המגדירים מלבן מינימלי העוטף את המתאר. תיבה זו שימושית למעקב, כי אפשר כך "לחתוך" את אזור היד בפריים או לעקוב אחר מיקומו בתמונה) לדוגמה, בהמשך מעדכנים את מסנן הקלמן לפי מיקום התיבה).
3. יחסי שטח (**Area Ratios**) נוסף לשטח המתאר) שטח היד בפועל (ושטח התיבה, $w \times h$) ניתן לחשב יחס בין שטח המתאר לבין שטח התיבה. ערך זה מעיד עד כמה הצורה "ממלאת" את התיבה, והוא עשוי לשמש לאבחנה בין מחוות) למשל, יד פתוחה תתפוס חלק ניכר מהתיבה, בעוד יד מקופלת או מחווה עם אצבע אחת בולטת עשויה להיות בעלת יחס שטח קטן יותר).
4. מרכז הכובד : לעיתים מבצעים חישוב רגעים גאומטריים כדי למצוא את מרכז הכובד של היד, או חישובים דוגמת היקף (Perimeter) ומומנטים מסדר גבוה לזיהוי צורות מורכבות יותר) למשל איתור מספר האצבעות על-ידי Convex Hull ו-(Convexity Defects).
- מסנן קלמן - (**Kalman Filter**) לא נלמד בהרצאות הקורס: נעשה שימוש במסנן קלמן כדי לשפר את יציבות המעקב אחר היד בין פריימים עוקבים. המסנן מסייע בהפחתת רעשים ומספק תחזיות מדויקות יותר למיקום היד, גם בתנאים של תנודות או רעשי רקע.
- שלבי היישום:
1. אתחול מסנן הקלמן: המסנן מאתחל עם מצב ראשוני המבוסס על מיקום וגודל היד בפריים הראשון. מטריצות המעבר, (Transition Matrix) מדידת הרעש, (Measurement Noise Covariance) ורעש התהליך (Process Noise Covariance) מוגדרות כדי לשקף את הדינמיקה הצפויה של תנועת היד.
 2. חיזוי (**Prediction**): בכל פריים, המסנן חוזה את המצב הבא של היד) מיקום וגודל תיבת החיבוק (בהתבסס על המצב הקודם ומודל התנועה הליניארי.

3. עדכון: **(Correction)** לאחר חיזוי המצב, מתקבלת מדידה חדשה של מיקום היד באמצעות זיהוי תווי המתאר. המסנן מעדכן את התחזית שלו על סמך המדידה החדשה, תוך שילוב בין התחזית למדידה בפועל, מה שמאפשר התמודדות עם רעשים ואי-ודאות.

שילוב ותרגום למחוות: לאחר איתור היד ועקיבה יציבה, מתבצע ניתוח של צורת היד כדי לזהות מחוות שונות, המיתרגמות לפקודות לשליטה ברחפן בזמן אמת.

שלבי התהליך:

1. איתור מתאר היד: **(Contour Detection)** לאחר יצירת מסכה בינארית המייצגת את היד, הקוד משתמש בפונקציה `cv2.findContours` כדי לאתר את המתארים בתמונה. המתאר הגדול ביותר נבחר כמתאר היד.

2. חישוב מעטפת קמורה: **(Convex Hull)** על מתאר היד, הקוד מחשב את המעטפת הקמורה באמצעות `cv2.convexHull`. המעטפת הקמורה היא הצורה הקמורה ביותר המכילה את כל נקודות המתאר, והיא מסייעת בזיהוי בליטות ושקעים ביד.

3. זיהוי פגמים קמורים: **(Convexity Defects)** הקוד מחשב את הפגמים הקמורים באמצעות `cv2.convexityDefects`, המזהים את האזורים בין המתאר למעטפת הקמורה. פגמים אלו מצביעים על שקעים בין האצבעות ויכולים לסייע בזיהוי מספר האצבעות המורמות.

4. זיהוי מחוות יד: הקוד מנתח את מספר הפגמים הקמורים ומאפיינים נוספים של המתאר כדי לזהות מחוות יד שונות. בהתאם למחוות המזוהות, הקוד מתרגם אותן לפקודות לשליטה ברחפן.

שלבי החישוב:

1. חישוב המומנטים

עבור מתאר נתון, מחושבים המומנטים הסטנדרטיים באמצעות הנוסחה:

$$M_{pq} = \sum_x \sum_y (x^p \cdot y^q \cdot I(x, y))$$

כאשר:

- כאשר $I(x, y)$ היא פונקציית האינדיקטור של הפיקסלים במתאר) כלומר 1, אם הפיקסל חלק מהמתאר, אחרת. 0).
- קו q -הן חזקות המגדירות את סדר המומנט.

2. חישוב קואורדינטות מרכז הכובד

חישוב הקואורדינטות של מרכז הכובד (x, y) באמצעות:

$$\bar{x} = \frac{M_{10}}{M_{00}}$$

$$\bar{y} = \frac{M_{01}}{M_{00}}$$

כאשר:

- M_{10} ו- M_{01} הם המומנטים הספציפיים מסדר ראשון
- M_{00} מייצג את שטח המתאר) המספר הכולל של הפיקסלים במתאר)

לסיום , תרגום המחוות לפקודות והאלגוריתם שיצרנו על מנת לקבל אותם:

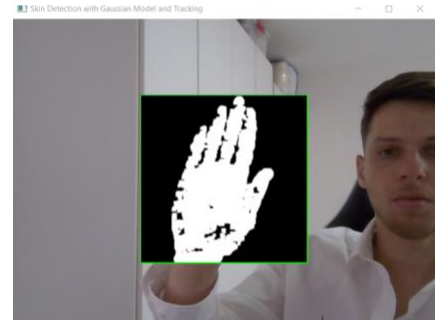
בעזרת החישובים המפורטים למעלה ניתן למצוא את מרכז המסה של הקונטור בנוסף ניתן למצוא את הנקודה הכי רחוקה על הקונטור ממרכז המסה ובכך ליצור וקטור היוצא ממרכז המסה עד לנקודה הכי רחוקה ממנו) במילים פשוטות אמצע היד זה מרכז המסה והאצבע היא הנקודה הרחוקה ביותר (ולכן ניתן לקבל את הכיוון של" למעלה , "למטה" , "ימינה" , "שמאלה" והרחפן יבצע את הפעולה.

1. אצבע אחת מורמת למעלה: **(UP)** הרפן יעלה למעלה .
 2. אצבע אחת מצביעה למטה: **(Down)** הרחפן מבצע נחיתה.
 3. אצבע מצביעה ימינה: **(Turn Right)** הרחפן מסתובב ימינה.
 4. אצבע מצביעה שמאלה: **(Turn Left)** הרחפן מסתובב שמאלה.
 5. שתי אצבעות מורמות למעלה: **(Take a Picture)** הרחפן מצלם תמונה. על מנת לבצע את פעולת התמונה הסתכלנו על ה 10% העליונים בקונטור וחיפשנו שם 2 קונטורים) כלומר חיפשנו 2 אצבעות (כפי שניתן לראות תנועת הצילום מתבצעת על ידי V עם האצבעות ולכן אנו מחפשים רק ב 10% העליונים של התמונה .
 6. יד פתוחה: **(Move Forward)** הרחפן נע קדימה. על מנת למצוא האם הפעולה היא להתקדם קדימה אנחנו מחשבים את היחס בין שטח היד) הקונטור (לבין המסגרת המקיפה אותו, ראינו במהלך הבדיקות כי היחס הזה נשמר בין כל הידים ולכן שמנו טווח שאם אנו נמצאים בו אז התנועה היא להתקדם קדימה .
- נציין ששמנו פונקציה שמחשבת בין ה 100 פריימים האחרונים ומעבירה את התנועה לרחפן רק במידה ויש אותה תנועה ב 80 מתוך ה 100 פריימים כלומר אנחנו עושים majority על מנת להבטיח שהתנועה המועברת לרחפן נכונה ב 80% לפחות .

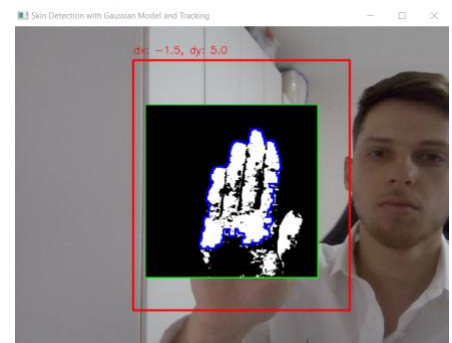


שלבי ביניים :

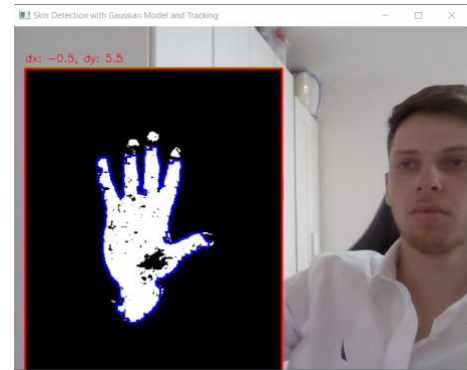
הקוד מבצע לימוד ועקיבה אחר כף היד של המשתמש בזמן אמת. בתחילת הריצה, המערכת לומדת את מאפייני הצבע של כף היד בתוך חלון בגודל מוגדר מראש, תוך שמירת דגימות צבע להמשך עיבוד.



לאחר שלב הלימוד, האלגוריתם מבצע סגמנטציה של כף היד על ידי סינון פיקסלים בגווי עור באמצעות מרחב HSV ופעולות מורפולוגיות להפחתת רעשים. לאחר מכן, נמצא הקונטור הגדול ביותר בתמונה, המייצג את כף היד, ומחושב Bounding Box לתחום אותה.



באמצעות מסנן קלמן, מתבצע מעקב אחר תנועת היד, תוך חישוב וקטור הכיוון בפריימים עוקבים. כך, האלגוריתם מאפשר זיהוי יציב ומדויק של תנועת היד גם בתנאים משתנים של תאורה ורקע.

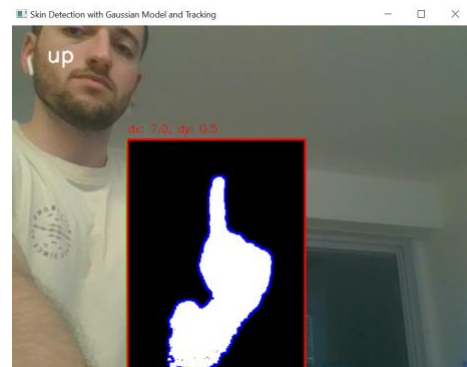


התוצאות הללו הגיעו לאחר שיחה עם תמי, שינוי שבו הכנסנו את הפעולות המורפולוגיות להפחתת רעשים ומסנן קלמן לפי עצתה.

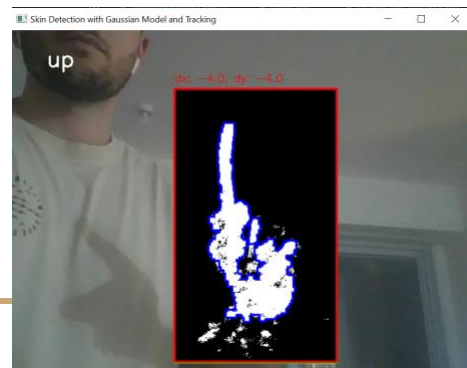
בחרנו לבדוק את ה confusion matrix על התנועה up על ידי הזזת היד על המסך לאורך 30 שניות שבהם 15 שניות אעשה את התנועה up ובשאר הזמן אעשה תנועות אחרות. עשיתי את הניסוי 3 פעמים כל פעם בתנאי תאורים שונים כפי שניתן לראות בתמונות המצורפות. דגמתי כל חצי שניה) סה"כ 60 דגימות (קיבלתי את התוצאות הבאות:

(1) תנאי אור טובים מאוד ניתן לראות שכמעט אין שגיאה כצפוי:

	Positive	Negative
True	30	3
False	0	27



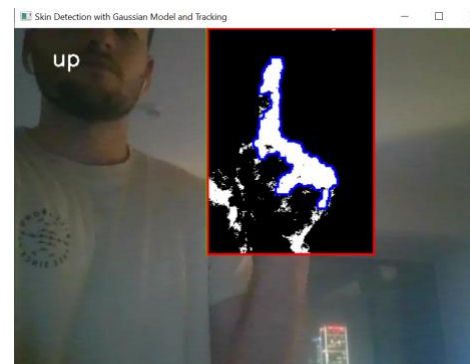
(2) תנאי אור פחות טובים:



	Positive	Negative
True	22	1
False	9	28

(3) תנאי אור מאוד נמוכים:

	Positive	Negative
True	15	15
False	16	14



סיכום:

בפרויקט שלנו, פיתחנו מערכת המאפשרת שליטה ברחפן באמצעות מחוות ידיים בזמן אמת, תוך שימוש בעיבוד תמונה מתקדם. המערכת מזהה את היד על ידי סינון צבע עור במרחב HSV מיישמת פעולות מורפולוגיות לשיפור הדיוק, ומאתרת את מתאר היד כדי לזהות מחוות שונות. לאחר מכן, מסנן קלמן משמש לייצוב המעקב אחר היד, מה שמאפשר שליטה מדויקת ואמינה ברחפן באמצעות מחוות יד טבעיות. במהלך הפיתוח, התמודדנו עם אתגרים בסביבת הספרייה, כגון הצורך בהוספת השהיות והגבלת תנועות הרחפן במרחק לכל פקודה, כדי להבטיח פעולה בטוחה ומדויקת. לאחר התמודדות עם אתגרים אלו, הצלחנו להשיג ניווט יציב של הרחפן באמצעות מחוות ידיים. ולבסוף הפרויקט עורר עניין בכנס עבור המבקרים, צוות הקורס ושאר הסטודנטים ונהננו מאוד להציגו.

ביבליוגרפיה:

- הרצאות הקורס - מבוא לעיבוד ספרתי של תמונה.
- הספרייה של הרחפן שאיתה התחברנו דרך API אליו בפיתוח tello :
