**המחלקה להנדסת תוכנה**

**פרויקט גמר – תשע"ז**

**זרימה אופטית בתמונות CT**

**Optical Flow on CT images.**

**מאת**

**שון אוטמזגין**

**מנחה אקדמי: דר' אסף שפיינר אישור: תאריך:**

**רכז הפרויקטים: דר' ראובן יגל אישור: תאריך:**

מערכות ניהול הפרויקט:

|  |  |  |
| --- | --- | --- |
| # | מערכת | מיקום |
| 1 | מאגר קוד | [GitHub](https://github.com/shon-otmazgin/Optical-Flow-on-CT-images.git) |
| 2 | יומן | [Google Calendar](https://calendar.google.com/calendar/embed?src=jdg4vvngq84llmrqm994it5ceo%40group.calendar.google.com&ctz=Asia/Jerusalem) |
| 3 | ניהול פרויקט | [GitHub Issues](https://github.com/shon-otmazgin/Optical-Flow-on-CT-images/issues) |
| 4 | הפצה |  |

# מילון מונחים, סימנים וקיצורים

* + - * **סירקות CT** - היא אחת מסוגי ההדמיות הרפואיות הנפוצות היא בטומוגרפיה ממוחשבת.

זוהי שיטת הדמיה (imaging) רפואית לא-פולשנית המשתמשת בטומוגרפיה הנוצרת על ידי עיבוד של מחשב, שנועד לתת תמונה תלת-ממדית של פנים הגוף.

* + - * **זיהוי קצוות (Edge Detection)** - הינה פעולה מתמטית מרחבית שמטרתה למצוא את קוי הקצה של עצמים בתמונה דיגיטאלית. קוי קצה (שפות) מוגדרים כגבול בין בין רמות בהירות שונות. בדרך כלל, השוני בין רמות בהירות שבינן עוברת שפה הוא חד. זיהוי קצוות הוא כלי בסיסי ב[עיבוד תמונה](https://he.wikipedia.org/wiki/%D7%A2%D7%99%D7%91%D7%95%D7%93_%D7%AA%D7%9E%D7%95%D7%A0%D7%94) ו[ראייה ממוחשבת](https://he.wikipedia.org/wiki/%D7%A8%D7%90%D7%99%D7%99%D7%94_%D7%9E%D7%9E%D7%95%D7%97%D7%A9%D7%91%D7%AA), המשמש בסיס לאלגוריתמים רבים לזיהוי עצמים וצורות בתמונות.
      * **סגמנטציה (Segmentation)** - ב[ראייה ממוחשבת](https://he.wikipedia.org/wiki/%D7%A8%D7%90%D7%99%D7%99%D7%94_%D7%9E%D7%9E%D7%95%D7%97%D7%A9%D7%91%D7%AA" \o "ראייה ממוחשבת) המונח מתייחס לתהליך של חלוקת [תמונה דיגיטלית](https://he.wikipedia.org/wiki/%D7%AA%D7%9E%D7%95%D7%A0%D7%94_%D7%93%D7%99%D7%92%D7%99%D7%98%D7%9C%D7%99%D7%AA) למקטעים (סגמנטים) כאוספים של [פיקסלים](https://he.wikipedia.org/wiki/%D7%A4%D7%99%D7%A7%D7%A1%D7%9C) הנמצאים זה ליד זה בתמונה. מטרת התהליך היא לפשט או לשנות את הייצוג של התמונה לאוסף של [אובייקטים](https://he.wikipedia.org/wiki/%D7%90%D7%95%D7%91%D7%99%D7%99%D7%A7%D7%98_(%D7%9E%D7%93%D7%A2%D7%99_%D7%94%D7%9E%D7%97%D7%A9%D7%91)) בעלי משמעות שניתן להתייחס אליהם לאחר מכן על ידי אלגוריתמים אחרים לראייה ממוחשבת או ל[עיבוד תמונה](https://he.wikipedia.org/wiki/%D7%A2%D7%99%D7%91%D7%95%D7%93_%D7%AA%D7%9E%D7%95%D7%A0%D7%94" \o "עיבוד תמונה), כגון מציאת גבולות או קווים. התוצאה של התהליך היא אוסף של מקטעים זרים המכסים את התמונה כולה.

ההפרדה נעשית על סמך תכונות של פיקסלים כגון [צבע](https://he.wikipedia.org/wiki/%D7%A6%D7%91%D7%A2), בהירות, [מרקם](https://he.wikipedia.org/wiki/%D7%9E%D7%A8%D7%A7%D7%9D) וכדומה.

* + - * **יער החלטות רנדולמי (Random decision forest)-** הינה שיטת למידה לסיווג ונסיגה של מידע, ע"י בניית המון עצי החלטות היוצרים ביחד יער.

השיטה מורכבת מ 2 שלבים:

1. שלב האימון – בו ישנו אוסף של מידע שאנחנו יודעים לגביו את סיווגו ובעזרתו אנו בונים את יער ההחלטות.
2. שלב הבדיקה – שבו "נכנס" ליער מידע שעוד לא פגשנו ובעזרת היער אנו מצפים שנקבל את הסיווג הנכון ביתר (האחוזים הגבוהים ביותר).

* **זרימה אופטית (Optical Flow)** – היא תבנית הנראית לעין של תנועה של אובייקטים, משטחים, וקצוות.

# מבוא

מאז שנות ה-70, עת החל השימוש בטומוגרפיה ממוחשבת(תמונות CT), היא הפכה לכלי חשוב בתחום [דימות רפואי](https://he.wikipedia.org/wiki/%D7%93%D7%99%D7%9E%D7%95%D7%AA_%D7%A8%D7%A4%D7%95%D7%90%D7%99) המשלים [צילום רנטגן רפואי](https://he.wikipedia.org/wiki/%D7%A7%D7%A8%D7%A0%D7%99_%D7%A8%D7%A0%D7%98%D7%92%D7%9F) ו[בדיקת אולטרה סאונד רפואית](https://he.wikipedia.org/wiki/%D7%91%D7%93%D7%99%D7%A7%D7%AA_%D7%90%D7%95%D7%9C%D7%98%D7%A8%D7%94_%D7%A1%D7%90%D7%95%D7%A0%D7%93_%D7%A8%D7%A4%D7%95%D7%90%D7%99%D7%AA).

בתחילת המאה ה-21 החל מתרחב השימוש בטומוגרפיה ממוחשבת אף ככלי בתחום [רפואה מונעת](https://he.wikipedia.org/wiki/%D7%A8%D7%A4%D7%95%D7%90%D7%94_%D7%9E%D7%95%D7%A0%D7%A2%D7%AA) ובדיקות תקופתיות לגילוי מחלות.

לטומוגרפיה ממוחשבת מספר יתרונות בתחום הדימות הרפואי על פני מכשירים המפיקים תמונות דו-ממדיות.

בתמונת טומוגרפיה ממוחשבת אין חפיפה של מבני הגוף זה על גבי זה בתמונה וניתן להתמקד באיבר המטרה.

מכשירי הטומוגרפיה הראשונים נזקקו למספר שעות כדי לבצע את הסריקה ולאסוף את הנתונים הגולמיים ולמספר ימים כדי לנתח את הנתונים.

כיום, סורקים מודרניים מסוגלים לסרוק תוך שניות ספורות, ולהפיק מודל תלת ממדי באופן כמעט מיידי. מהירות זו מאפשרת לסרוק חלקי גוף נעים כגון ריאות שנושמות או לב שפועם, ולכן עולה חשיבותו של הרופא הרדיולוג, אשר קובע את איכות ביצוע הבדיקות ובאמצעות הידע הקליני שלו מזהה את המחלה וחומרתה.

# תיאור הבעיה

רוב הציבור אינו מכיר את מקצוע הדימות (רדיולוגיה) ואינו מודע מספיק לחשיבות הרבה של איכות הביצוע של בדיקת הסיטי ולחשיבותו הגדלה והולכת של הרופא המפענח- הרדיולוג.

פענוח הבדיקה מורכב ודורש זמן ארוך.

בנוסף לכך, פענוח הבדיקות אינו מבוצע עוד רק במישור אחד, כפרוסות, אלא יש אפשרות להכנת שחזורים דו ותלת מימדים אשר מעלים את דיוק הבדיקה במקרים רבים. הכנה של שחזורים מתקדמים לפני ניתוחים אונקולוגים מורכבים ע"י הרדיולוג, מסיעת מאוד לכירורג בהחלטה על התאמת החולה לניתוח ויכולה לשנות את הגישה הניתוחית.

בדיקה איכותית ופענוח מפורט ומקצועי, תלויים ברדיולוג מיומן ומנוסה והם קריטים לבריאות החולה.

הידע הקליני של הרדיולוג, ניסיונו ואיכות הפענוח שהוא מספק יאפשרו אבחון מדויק של מחלות, מציאת גידולים ממאירים, הערכה נכונה של התגובה לטיפולים אונקולוגים (כימותרפיה והקרנות).

ניתן לדמות את הרדיולוג לקצין מודיעין אשר מכין תרשימים, ומפות למבצע מיוחד.

ככל שרמת התרשימים תהייה מדויקת יותר, ופענוח המצב מדויק יותר, כך סיכויי ההצלחה של "הלוחם", הרופא המטפל או המנתח, יהיו **גבוהים** יותר בקרב על חיי החולה.

## דרישות ואפיון הבעיה

בעקבות חשיבות וקריטיות איכות פענוח הדימות ע"י הרדיולוג, נרצה לעזור לפענוח הדימות מבחינה טכנולוגית בכך שנעקוב באופן מסומן אחר התפתחות איברים מסויימים(לפי דרישה) בסריקה.

הדרישות:

1. הקלט יהיה תמונת הCT הראשונה ובה סימון של האיבור אחריו נרצה לבצע את המעקב בשאר בתמונות.
2. יצירת סגמנטציה בעזרת קצוות(Edges), וגבולות מדויקים ככל הניתן בין האיבר אחריו אנו עוקבים לבין שאר אובייקטים/איברים בסריקה.
3. הגדלת המעקב אחר כמה איברים בסריקה.

## הבעיה מבחינת הנדסת תוכנה

Optical Flow אינו תהליך חדש, ומשתמשים בו בהמון תחומים הקשורים לעיבוד תמונה.

ההבדל באלגוריתמים השונים הוא רמת הדיוק ויכולת הדיוק.

האתגר בפרוייקט הוא לקחת כלי שעובד בתחומים שונים בעיבוד התמונה כגון תמונות צבעוניות, תמונות נוף ובכלל תמונות שניתן להפיק מהם הרבה מידע ולשלבו בתחום של תמונות רפאויות ובמקרה הפרטי שלנו תמונות CT.

# תיאור הפתרון

הפתרון שאנו מציעים בפרוייקט זה לבעיה הנתונה היא יצירת סגמנטציה אוטומטית לאיבר מסוים בסריקת CT ובכך יהיה ניתן לזהות בצורה יותר חד משמעית את םיקסלים מוסיימים בסריקה שייכים לאיבר אותו בודקים או לא.

## מהי המערכת

אנו נשתמש באלגוריתמים שונים, נבחן את השוני בינהם וניקח את האלגוריתם עם התוצאה הטובה ביותר עבורנו.

1. TOOLBOX יודעים בסביבת MATLAB.
2. Epic Flow – **נספחים איור 1.1**

## תהליכים ונתוני המערכת

קלט המערכת: סריקת CT וסגמנטציה בודדת של האיבר אחריו נרצה "לעקוב".

פלט המערכת: סגמנטציה אוטומטית של האיבר בתמונות הבאות.

## תיאור הפתרון המוצע

הפתרון ימומש בעזרת טכניקות ואלגוריתמי Optical flow לקבלת זרימה של אובייקט על רצף של תמונות.

בהינתן סגמנטציה בודדת של איבר בסריקת CT, נרצה לקבל את הסגמנטציה אוטומטית בתמונה הבאה וכן הלאה עד סוף הסריקה (את הזרימה), ובכך נוכל לעקוב ולפענח בצורה מדויקת יותר איברים ספציפים בסריקה – **נספחים איור 1.2**

## תיאור הכלים המשמשים לפתרון

1. [MATLAB OPTICAL FLOW TOOLBOX](http://www.mathworks.com/matlabcentral/fileexchange/44400-tutorial-and-toolbox-on-real-time-optical-flow)
2. ITK – SNAP - מכילה אופציות לסגמנטציה אינטראקטיבית ותצוגה.

# סקירת עבודות דומות בספרות והשוואה

האב טיפוס של הפרוייקט יכלול מימוש והטמעה של פרוייקט מחקרי של קבוצת [Inria](http://www.inria.fr/en/centre/grenoble) וקבוצת  [Laboratoire Jean Kuntzmann](http://www-ljk.imag.fr/ljk_en.html)הנקרא EpicFlow.

במחקר זה שולבו טכניקות שונות ליצירת אלגוריתם חדש שבו יש רמת דיוק גבוה של זרימת אובייקט על רצף תמונות.

המחקר הינו חדש(2015) והוא ממחיש שילוב של 2 אלגוריתמים שונים (אחד מהם של מיקרוסופט) היוצרים ביחד את ה Optical Flow על אובייקטים(ראה נספח מאמרים).

אנו נשתמש במחקר זה, עם שוניים בסוג הנתונים – כלומר במקום תמונות נוף נשתמש בתמונות רפואיות.

# נספחים

ספרות, תרשימים נוספים, תכנון הפרויקט, טבלת ניהול סיכונים, טבלת דרישות (URD),

## רשימת ספרות \ ביבליוגרפיה

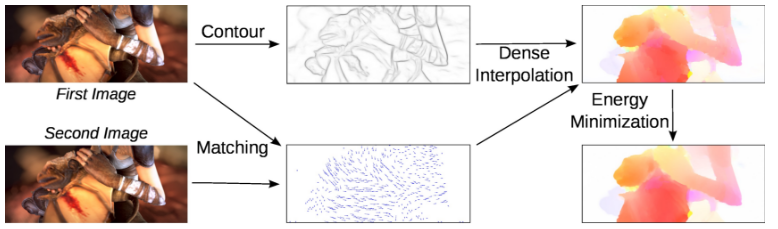
1. מאמרים:

* [EpicFlow: Edge-Preserving Interpolation of Correspondences for Optical Flow](http://lear.inrialpes.fr/src/epicflow/)
* [Structured Forests for Fast Edge Detection](https://www.microsoft.com/en-us/research/publication/structured-forests-for-fast-edge-detection/)
* [DeepMatching: Deep Convolutional Matching](http://lear.inrialpes.fr/src/deepmatching/)
* [Decision forests a unified framework for classification](https://www.microsoft.com/en-us/research/publication/decision-forests-a-unified-framework-for-classification-regression-density-estimation-manifold-learning-and-semi-supervised-learning/)

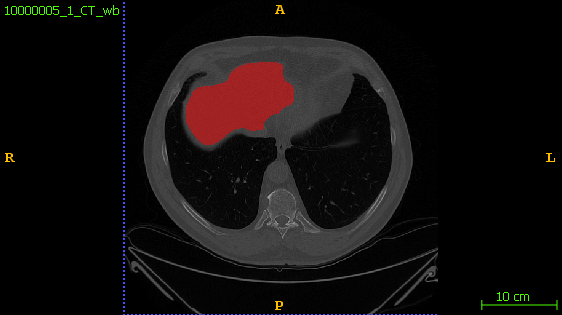
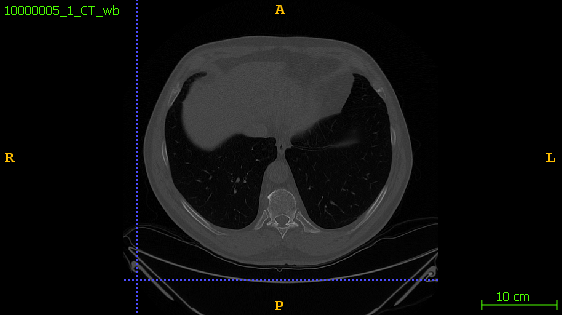
1. הגדרות ומושגים:

* [Edge detection](https://en.wikipedia.org/wiki/Edge_detection)
* [Image segmentation](https://en.wikipedia.org/wiki/Image_segmentation)
* [Random decision forest](https://en.wikipedia.org/wiki/Random_forest)
* [Optical Flow](https://en.wikipedia.org/wiki/Optical_flow)

## תרשימים וטבלאות



**איור מספר 1.1:** מתאר את האלגוריתם המחקרי של מיקרוסופט למציאת Optical Flow



איור מספר 1.2: מתאר את המצב אליו נרצה להגיע – מעקב וסימון של איבר בסריקת CT.

## תכנון הפרויקט

|  |  |
| --- | --- |
| 27.7.16 | פגישת היכרות עם אסף למטרות יעוץ. |
| 30.7.16 | סיכום נושא פרוייקט עם המנחה. |
| 15.8.16 | קבלת חומר לקריאה מהמנחה. |
| 6.9.16 | התחלת קריאה ולמידה על תמונות רפואיות |
| 7.9.16 | המשך למידת תמונות רפואיות |
| 10.9.16 | המשך למידה. |
| 11.9.16 | המשך למידה. |
| 12.9.16 | המשך למידה. |
| 9.10.16 | המשך למידה. |
| 16.10.16 | המשך למידה. |
| 23.10.16 | המשך למידה. |
| 30.10.16 | המשך למידה. |
| 7.11.16 | פגישת סטאטוס עם המנחה. |
| 20.11.16 | הדרכה מהמהנחה על עבודה עם תמונות CT. |
| 4.12.16 | הגשת הצעת פרויקט למנחה + פגישה עם המנחה. |
| 11.12.16 | המשך עבודה ומחקר. |
| 20.12.16 | המשך עבודה ומחקר. |
| 27.12.16 | המשך עבודה ומחקר. |
| 30.12.16 | המשך עבודה ומחקר. |
| 1.1.17 | המשך עבודה ומחקר. |
| 23.1.17 | המשך עבודה ומחקר. |
| 30.1.17 | המשך עבודה ומחקר. |
| 12.2.17 | הגשת אב טיפוס |
| 22.2.17 | המשך עבודה ומחקר. |
| 1.3.17 | המשך עבודה ומחקר. |
| 8.3.17 | המשך עבודה ומחקר. |
| 29.3.17 | סטאטוס פרויקט וקביעת מטרות לסוף הפרויקט |
|  |  |

## טבלת סיכונים

|  |  |  |  |
| --- | --- | --- | --- |
| **#** | **הסיכון** | **חומרה** | **מענה אפשרי** |
| 1 | הערכת לוחות זמני פיתוח שגויים | 5 | הערכה מחודשת של לוחות הזמנים. עדכון רשימת הדרישות על פי עדיפות - מהגבוה לנמוך.  במקרים קיצוניים, נוותר על דרישות שאינן דרישות מפתח. |
| 2 | התממשקות לתוכנות חיצוניות שאינן מספקות API | 2 | מציאת תוכנה שכן מספקת API בסיסי. |
| 3 | הערכה שגויה של גודל המערכת | 2 | קריאה מקיפה של מסמכי המחקר של מיקרוסופט |
| 4 | קריאת תצלומי CT | 2 | למידה מקדימה של הנושא |
| 5 | טכניקות לקויות ב MATLAB | 3 | לקיחת קורסים ONLINE |
| 6 | טכניקות לקויות ב C++ | 2 | לקיחת קורסים ONLINE |
| 7 | ביצועי זמן אמת בלתי מספקים | 4 | הערכה מחדש של הקוד הקיים עם אפשרות ל-refactory. מימוש אלגוריתם יעיל יותר. |
|  |  |  |  |

## רשימת\טבלת דרישות

**טבלת דרישות (User Requirement Document)**

|  |  |
| --- | --- |
| מס' דרישה | תיאור |
| 1 | קלט תמונות סגמנטציה של איבר אחריו נרצה "לעקוב". |
| 2 | התאמת האלגוריתם הממש OPTICAL FLOW |
| 3 | יצירת סגמנטציה בתמונה הבאה של הסריקה וכן האלה. |
| 4 | הגדלת מספר האיברים אחריו עוקבים עד 4 |
| 5 | דיוק המערכת ב 90 אחוז. |