

דו"ח סיכום פרויקט: ב'

# [סיווג מחלות לב] [Classification of Valvular Heart Diseases]

מבצעים:

[Itay Damary]

[איתי דמארי]

[Etay Sela]

[איתי סלע]

מנחים:

[Hadas Ofir]

[הדס אופיר]

[Yehonatan-Itay Segman]

[יהונתן-איתי סגמן]

סמסטר רישום: אביב תשפ"ג

תאריך הגשה: ספטמבר, 2023

בשיתוף עם: [Sanolla] 

**P** 1234-2-19

## תודות

ברצונם של כותבי ספר זה להודות למעבדת SIPL בכלל ולמנחים הרשומים מעלה בפרט על הנחיית פרויקט זה. התמיכה והעזרה בליווי הפרויקט תרמו רבות הן להצלחתו והן להכשרתם הלימודית של הכותבים. נוסף על כך, נודה לחברת Sanolla על ההזדמנות לעבוד על פרויקט מרתק ומסקרן בעל יישומים ממשיים בעולם הרפואה.

## תוכן עניינים

2	תודות	
6	1. תקציר	
7	Abstract	1.
8	2. מבוא	2.
8	2.1 מוטיבציה	2.1
8	2.2 עיקרי הדברים	2.2
8	2.3 אתגרים מרכזיים	2.3
9	3. סקר ספרות	3.
9	3.1 קולות לב	3.1
10	3.2 סגמנטציה	3.2
10	3.2.1 שיטות מעטפת	3.2.1
11	3.2.2 שיטות הסתברותיות	3.2.2
12	3.3 סיווג	3.3
12	3.3.1 מסווגים בסיסיים	3.3.1
12	3.3.2 רשתות עמוקות	3.3.2
13	4. בעיית הסגמנטציה	4.
13	4.1 שיטת מעטפת	4.1
14	4.1.1 אלגוריתם	4.1.1
16	4.1.2 תוצאות	4.1.2
19	4.2 שיטת STFT	4.2
19	4.2.1 אלגוריתם	4.2.1
22	4.2.2 תוצאות	4.2.2
26	5. בעיית הסיווג	5.
26	5.1 פיצ'רים תיאוריים ו-SVM	5.1
26	5.1.1 חילוף פיצ'רים	5.1.1
27	5.1.2 סיווג	5.1.2
28	5.2 מקדמי MFC ועצי החלטה	5.2
28	5.2.1 חילוף פיצ'רים	5.2.1
31	5.2.2 סיווג	5.2.2
31	5.3 ספקטרוגרמת MFC ורשת נוירונים	5.3
31	5.3.1 חילוף פיצ'רים	5.3.1
32	5.3.2 סיווג	5.3.2
33	5.4 תוצאות	5.4
33	5.4.1 חיזוי בהינתן מחזור בודד	5.4.1
34	5.4.2 חיזוי בהינתן הקלטה	5.4.2
34	5.4.3 הרחבה – אימון מודל ה-ResNet	5.4.3
35	5.4.4 הרחבה – הפרדה ע"י פיצ'רים תיאוריים	5.4.4
37	6. מסקנות	6.
38	7. סיכום	7.
39	8. עבודה עתידית	8.
40	9. רשימת מקורות	9.

## רשימת איורים

9	איור 3.1, דוגמה לאות לב
10	איור 3.2, מחזורי לב אופייניים של תסמונות נפוצות
11	איור 3.3, דוגמה לחישוב מעטפת מתוך אות לב
11	איור 3.4, דוגמה לשרשרת מרקוב לביצוע סגמנטציה של קולות לב
14	איור 4.1, Envelope Segmentation blocks diagram
17	איור 4.2, שחזור פסגה באמצעות חיתוך סף אדפטיבי
17	איור 4.3, הורדת split noises
18	איור 4.4, הורדת פסגות כפולות
18	איור 4.5, תיוג פסגות בשיטת מעטפת
19	איור 4.6, טעות בתיוג פסגות בשיטת מעטפת
20	איור 4.7, STFT Segmentation block diagram
22	איור 4.8, הערכת ביצועים להערכת קצב לב, trainHR
23	איור 4.9, הערכת ביצועים להערכת קצב לב, Rest of sanolla_v2 dataset
24	איור 4.10, דוגמה להערכת קצב הלב
24	איור 4.11, דוגמה לסגמנטציה טובה
25	איור 4.12, דוגמה לסגמנטציה לא טובה
26	איור 5.1, דיאגרמת בלוקים של מסווג SVM
28	איור 5.2, דיאגרמת בלוקים של מסווג DT
30	איור 5.3, מסנני מל
31	איור 5.4, ספקטרוגרמת MFC
31	איור 5.5, דיאגרמת בלוקים של מסווג ResNet
32	איור 5.6, ארכיטקטורת רשת ResNet50
33	איור 5.7, Residual Block
35	איור 5.8, ResNet Overfitting
35	איור 5.9, הפרדה ע"י פיצורים תיאוריים

## רשימת טבלאות

9	טבלה 3.1, השוואה בין רעשי S1,S2
13	טבלה 4.1, מאגרי נתונים לשלב הסיווג
17	טבלה 4.2, הערכת ביצועי אלגוריתם הערכת דופק הלב
22	טבלה 4.3, הערכת ביצועי אלגוריתם הערכת דופק הלב
23	טבלה 4.4, הערכת ביצועי אלגוריתם הסגמנטציה על מאגר Sanolla_v2
24	טבלה 4.5, פילוח טעויות סגמנטציה
26	טבלה 5.1, מאגרי נתונים לשלב הסיווג
34	טבלה 5.2, סיכום תוצאות הסיווג בהינתן מחזור בודד
34	טבלה 5.3, סיכום תוצאות הסיווג בהינתן הקלטה

ספר זה מתאר את פרויקט "Classification of Valvular Heart Diseases" שהתבצע במסגרת פרויקט לימודי בהנחיית מעבדת SIPL בטכניון ובשיתוף עם חברת Sanolla. מטרת הפרויקט היא בניית מסווג בינארי המבחין בין הקלטת רעשי לב של אדם בריא לכזו של אדם חולה עבור הקלטה נתונה הנלקחת ממכשיר Voqx - סטטוסקופ חכם של חברת Sanolla. מערכת מסוג זה תאפשר לחסוך זמן של רופאים מומחים הנדרשים לבצע אוסקלציה ידנית, תנגיש יכולת לאבחון ראשוני של מחלות לב לכלל הציבור ותמנע טעויות אנוש. במסגרת ספר זה מתוארים הפתרונות הקיימים היום לבעיה בספרות המקצועית, הפתרון הנבחר בחלוקה לשני שלבים (סגמנטציה וסיווג), ותוצאות האלגוריתם בהשוואה לסטנדרט הקיים. נבחנו מספר פתרונות שונים לכל אחד משלבי הבעיה כאשר הפתרון הטוב ביותר הגיע ל-80% דיוק בסיווג.

## 1. Abstract

This book describes the "Classification of Valvular Heart Diseases" project that was carried out as part of a study project led by the SIPL laboratory at the Technion and in collaboration with the "Sanolla" Company. The purpose of the project is to build a binary classifier that distinguishes between a recording of heart sounds of a healthy person and that of a sick person for a given recording taken from a Voqx device - a smart stethoscope from the Sanolla Company. A system of this type will save the time of specialist doctors who are required to perform manual auscultation, will make the ability for initial diagnosis of heart diseases accessible to the general public and will prevent human errors. As part of this book, the solutions that exist today to the problem in the professional literature are described, the chosen solution divided into two stages (segmentation and classification) are explained, and the results of the algorithm compared to the existing standard is presented. Several different solutions were tested for each of the stages of the problem, with the best solution reaching 80% classification accuracy.

## 2. מבוא

### 2.1 מוטיבציה

מערכת המסווגת ברמת דיוק גבוהה בין הקלטות רעשי לב של אדם בריא לכאלו של אדם חולה תאפשר לחסוך זמן של רופאים מומחים הנדרשים לבצע אוסקלציה ידנית, תנגיש יכולת לאבחון ראשוני של מחלות לב לכלל הציבור ותמנע טעויות אנוש.

### 2.2 עיקרי הדברים

הבעיה חולקה לשניים: סגמנטציה של אותות הלב למחזורים בודדים, וסיווג כל מחזור בנפרד. עבור שלב הסגמנטציה נבחנו שני פתרונות, שיטת מעטפת ושיטת  $STFT$  כאשר השנייה הראתה יתרונות משמעותיים על פני הראשונה ועל כך יורחב בפרק 4. במסגרת שלב הסיווג נבחנו שלושה סוגי מסווגים שונים, כל אחד מהם ניזון מפיצ'רים שונים אשר מחולצים ממחזור הלב. המסווג אשר הראה את התוצאות המבטיחות ביותר הוא מסווג מסוג  $ResNet$  אשר ניזון מ- $MFCC$  dynamic features של המחזור ועל כך יורחב בפרק 5.

### 2.3 אתגרים מרכזיים

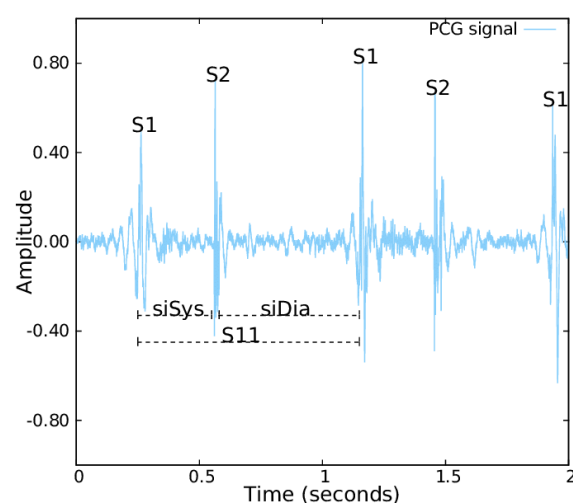
בניית המערכת לסיווג אותות לב נתקלה באתגרים הן ביולוגיים והן טכנולוגיים. אותות הלב שונים זה מזה כתלות בתסמונות השונות וקצבי הלב האינדיבידואליים. השוני מתבטא לעיתים אף בהעלמות חלק מרעשי הלב ועל כך יורחב בחלק 3.1 ובחלק 4.1.2.



### 3. סקר ספרות

#### 3.1 קולות לב

מחזור הלב מחולק לשני חלקים עיקריים: סיסטולי ודיאסטולי. החלק הסיסטולי בד"כ קצר יותר מהחלק הדיאסטולי. נקודת ההתחלה של כל אחד מהחלקים, הסיסטולי והדיאסטולי, הינה רעש דומיננטי,  $S1$  ו- $S2$  בהתאמה. באיור 3.1 מוצג אות לב לדוגמה כאשר על גבי אות הלב מסומנים רעשי  $S1, S2$  והחלקים הסיסטולי והדיאסטולי.



איור 3.1, דוגמה לאות לב

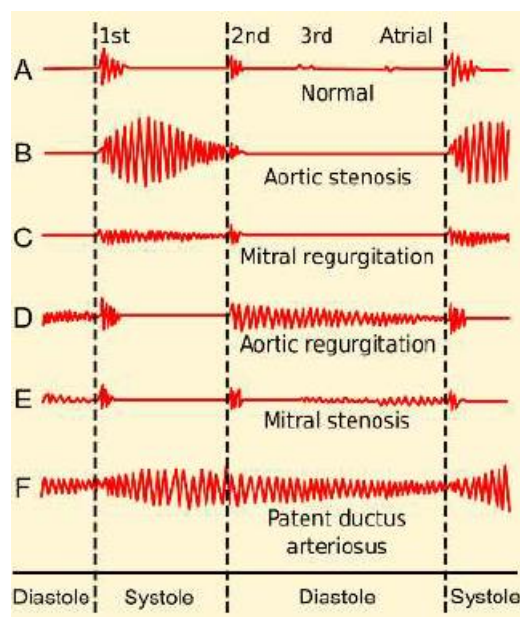
באיור 3.1, ציר x הוא ציר הזמן ביחידות של [s] וציר y מייצג את משרעת אות הלב ביחידות שרירותיות.

רעשי  $S1$  ו- $S2$  נובעים מסגירת ה-*atrioventricular valves* וה-*semilunar valves* בהתאמה. רעשים אלו נבדלים ביניהם הן במשכם הזמני והן בחתימתם התדרית. בעוד שרעש  $S1$  הוא עמום וממושך, רעש  $S2$  הוא חד וקצר. טבלה 3.1 משווה בין שני הרעשים.

Heart Sound	Frequency Range [Hz]	Duration [s]
$S1$	10 – 200	0.12 – 0.15
$S2$	20 – 250	0.08 – 0.12

טבלה 3.1, השוואה בין רעשי  $S1, S2$

בקולות לב לא-נורמליים (*Abnormal Heart Sounds*) עשויים להופיע רעשים נוספים ואף רעש  $S2$  עשוי "להיבלע" כלומר להעלם מהמחזור. איור 3.2 מציג מחזור לב אופייני של כמה מהתסמונות הנפוצות. הפרויקט הנוכחי איננו מתרכז בהבדלה בין תסמונות שונות אלא אך ורק בהבחנה בין אות נורמלי ללא-נורמלי.



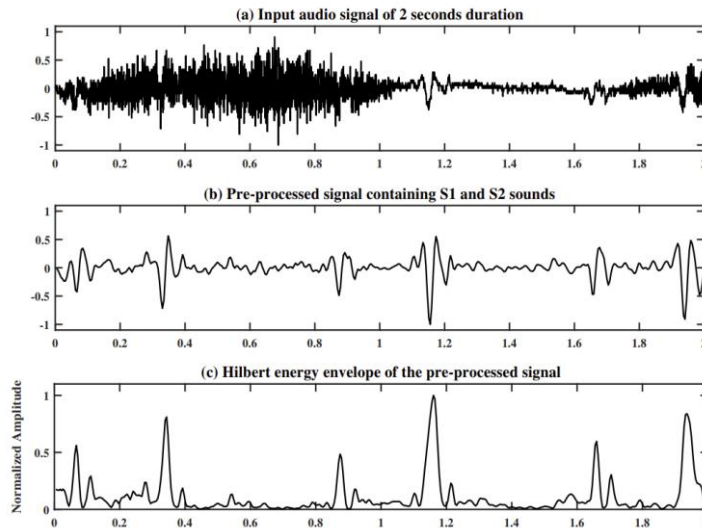
איור 3.2, מחזורי לב אופייניים של תסמונות נפוצות

### 3.2 סגמנטציה

מטרת הסגמנטציה היא חלוקת אות לב למחזורים השונים ולעיתים גם חלוקה פנים מחזורית לחלקים סיסטולי ודיאסטולי. נכון לכתיבת ספר זה, בספרות המקצועית קיימים שני פתרונות עיקריים לבעיה.

#### 3.2.1 שיטות מעטפת

שימוש בכלים מתמטיים לטובת ביצוע המשימה בשני שלבים. בשלב הראשון תחושב מעטפת האות, למשל ע"י חישוב התמרת הילברט או אנטרופיית שאנון. בשלב השני, באמצעות חיתוכי סף על גבי המעטפת ימצאו רעשי הלב הרלוונטיים לטובת ביצוע הסגמנטציה. איור לדוגמא המציג את חיתוך הסף שביצעו *H. Liang et al.* על גבי מעטפת מבוססת התמרת הילברט מופיע באיור 3.3.

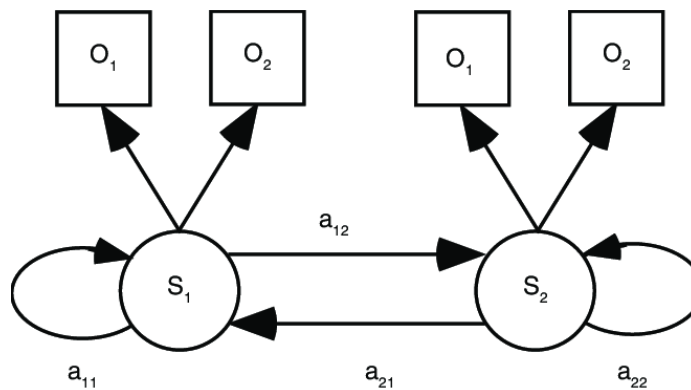


איור 3.3, דוגמא לחישוב מעטפת מתוך אות לב

באיור 3.3 ישנם 3 גרפים. למעלה, אות הכניסה. באמצע, אות לאחר עיבוד מקדים וסינון רעשים. למטה, מעטפת הילברט של האות המעובד. ציר x הוא ציר הזמן ביחידות של [s] וציר y מייצג את משרעת מעטפת האות ביחידות שרירותיות. החסרונות של שיטות מעטפת היא התמודדות עם רעשים. בחישובי המעטפת שיאים בעלי משמעות עלולים ללכת לאיבוד כתוצאה מהיבלעותם לתוך רעשי הסביבה.

### 3.2.2 שיטות הסתברותיות

שיטות הסתברותיות הוצעו על מנת להתגבר על חסרונותיהם של שיטות המעטפת. מטרתן של שיטות הסתברותיות היא לאפיין את רעשי הלב באמצעות פיצורים דיסקרמנטיביים ובתוכם הבעיה מבוסס על שרשראות מרקוב ונקרא *HMM – Hidden markov model*. בגישה זו, המודל מסווג את הסיגנל באמצעות שרשרת מרקוב אשר מצביה מוגדרים לפי מצבי מחזור הלב והסתברויות המעבר בה מתפתחות עם הזמן. איור 3.4 מציג דוגמא לשרשרת כזו.



איור 3.4, דוגמא לשרשרת מרקוב לביצוע סגמנטציה של קולות לב

החיסרון העיקרי של גישה זו הוא שהפיצ'רים המשמשים לאפיון הרעשים הינם בעלי *variance* גבוה כתלות בפרמטרים כגון גיל הפציינט ואבחנתו.

### 3.3 סיווג

מטרת הסיווג בפרויקט זה היא בניית סוכן המקבל מחזור לב בודד ומסווגו לאחת משתי מחלקות: נורמלי, לא-נורמלי. נציין כי רוב הפתרונות הקיימים היום מתייחסים לסיווג של הקלטה במלואה (בניגוד לסיווג מחזור בודד). נכון לכתיבת ספר זה, בספרות המקצועית קיימים שני פתרונות עיקריים לבעיה.

#### 3.3.1 מסווגים בסיסיים

מסווגים בסיסיים מתבססים על פי רוב על חילוץ של מספר סופי ומצומצם של פיצ'רים מתוך ההקלטה. הפיצ'רים נבחרים ידנית ומותאמים לצרכי המערכת. לאחר חילוץ הפיצ'רים נעשה שימוש במסווג כגון *Decision Tree*, *SVM* או *KNN* על מנת להפריד בין המחלקות השונות.

#### 3.3.2 רשתות עמוקות

רשתות עמוקות מסתמכות על מפות פיצ'רים רחבות יותר ממסווגים בסיסיים. פיצ'רים מסוג זה בד"כ מחולצים ע"י הסתכלות משולבת על האות במישור הזמן ומישור התדר באמצעות התמרות כמו *STFT*, *DWT*, *MFC* ועוד.

#### 4. בעיית הסגמנטציה

הקלט של בעיית הסגמנטציה הוא קובץ *wav*. אשר מכיל הקלטה של אות לב. במקרים בהם עבדנו עם מידע גולמי של חברת *Sanolla* (הרחבה בהמשך), אות הקלט הינו תוצאת החיסור בין האות הנקלט על ידי המיקרופון שמכוון כלפי פנים הגוף לבין המיקרופון שמכוון כלפי חוץ. פלט שלב זה הוא רשימה של זמנים בהם נפתח מחזור לב חדש בהקלטה הנתונה.

עבור בעיית הסגמנטציה נבחנו 2 פתרונות שונים. ראשית, נבחן פתרון בשיטת מעטפת אשר מומש באמצעות מעטפת הילברט ומספר שלבים של חיתוכי סף (4.1). פתרון זה נזנח עקב תוצאות לא מספקות וראה הרחבה ב-4.1.2. הפתרון השני שנבחן הוא פתרון בשיטת *STFT* במסגרתה הוערך דופק ההקלטה ועובדה התמרת ה-*STFT* לצורך תיוג רעשי *S1* (4.2).

כל אחת מהשיטות מומשו והוערכו על מספר מאגרי נתונים. מאגר הנתונים הראשון שנבחן, *peterjbentley*, הוא מאגר פתוח שנלקח מתוך אתגר אינטרנטי לסיווג קולות לב. רוב ההקלטות במאגר נלקחו באמצעות אפליקציית *iStethoscope* וחלקן הקטן באמצעות סטטוסקופ דיגיטלי. מיקומי רעשי *S1* וגם רעשי *S2* מתויגים רק עבור חלק מהמאגר המכיל הקלטות של קולות נורמליים בלבד. מאגר הנתונים השני עליו נבחן האלגוריתם הוא מאגר *sanolla\_v2* אשר הוקלט באמצעות מכשיר *Voqx* מבית חברת *Sanolla*. מיקומי רעשי *S1* בלבד מסומנים עבור כל ההקלטות במאגר. טבלה 4.1 משווה בין מאגרי הנתונים השונים ששימשו את הפרויקט לשלב זה.

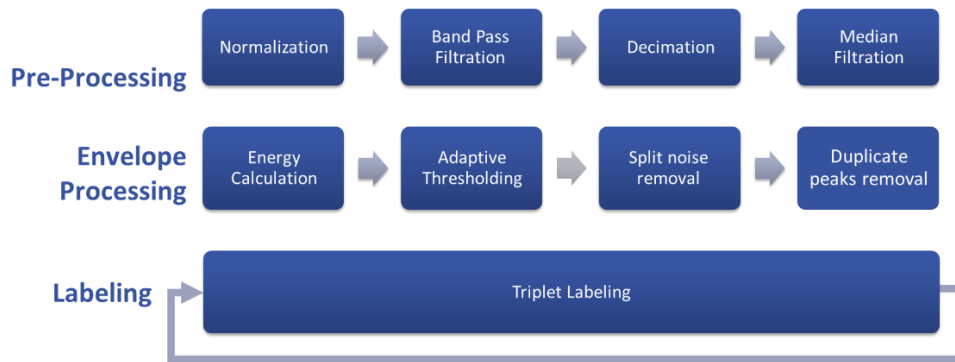
מאגר	<i>peterjbentley</i>	<i>sanolla_v2</i>
מקור	אתגר אינטרנטי	חברת <i>Sanolla</i>
מכשיר הקלטה	<i>iStethoscope</i>	<i>Voqx</i>
מס' הקלטות	312	318
מס' הקלטות סימפטומטיות	112	11
תיוג <i>S1</i>	חלקי, עבור קולות נורמליים בלבד	קיים
תיוג <i>S2</i>	חלקי, עבור קולות נורמליים בלבד	לא קיים

טבלה 4.1, מאגרי נתונים לשלב הסיווג

#### 4.1 שיטת מעטפת

שיטה זו היא שיטת הסגמנטציה הראשונה שנבחנה והיא מתבססת על חישובי סף על גבי מעטפת האות הנתון. שיטה זו נבחרה לאור יתרונותיה על השיטות האחרות מבחינת פשטות הפתרון ואובסרוויליות (*observability*). השיטה נזנחה לאור תוצאות לא מספקות וראה הרחבה ב-4.1.2.

דיאגרמת בלוקים כללית של שלב הסגמנטציה בשיטת מעטפת מוצגת באיור 4.1 :



איור 4.1, Envelope Segmentation blocks diagram

#### שלב ראשון – עיבוד מקדים

בשלב הראשון של הפתרון שמימשנו מתבצע עיבוד מקדים להקלטה לפי השלבים הבאים :

1. ביצוע נורמליזציה לכל הדגימות של המידע המתקבל מההקלטה כך שמתקבל אות בתחום  $[-1,1]$ .
2. סינון האות באמצעות מסנן 8<sup>th</sup> order Butterworth filter בתחום התדרים  $[25, 400][Hz]$ . תחום זה נבחר מאחר שבסקר הספרות עלה שזהו תחום התדרים שמרכיבים את קולות הלב, ובנוסף מאפשר סינון של רעשי רקע לא רצויים שעלולים להפריע כמו קולות הנשימה המופקים בריאות ומקורם בתדרים גבוהים יותר.
3. דצימציה לתדר של  $f_s = 900[Hz]$  למניעת תופעת התחזות.
4. העברת הסיגנל במסנן חציון עם גרעין בגודל 7 דגימות, כ- $8[ms]$ , לצורך סינון נוסף של רעשים. הפתרון בו מתבצעים חישובי התמרות לצורך מעבר לפתרון במישור התדר.

#### שלב שני – חישוב ועיבוד מעטפת

1. עבור האות שעבר את העיבוד המקדים מחושב האות האנליטי המתואר ע"י נוסחה 4.1 :

$$x_a = F^{-1}(F(x)2U) = x + iy$$

נוסחה 4.1, חישוב האות האנליטי

כאשר  $F$  היא התמרת פורייה,  $U$  היא פונקציית מדרגה ו- $y$  היא התמרת הילברט של האות  $x$  אשר ניתנת לכתיבה ע"י נוסחה 4.2 :

$$H(x(t)) = \frac{2}{\pi} \lim_{\epsilon \rightarrow 0} \int_{\epsilon}^{\infty} \frac{x(t-\tau) - x(t+\tau)}{2\tau} d\tau$$

נוסחה 4.2, חישוב התמרת הילברט

כאשר  $H$  היא התמרת הילברט של האות  $x$ .

2. כעת מתבצע חיתוך סף אדפטיבי על תוצאת המעטפת:

א. מחושב מספר הפסגות המינימלי  $n_{mini}$  כתלות באורך ההקלטה  $TR$  ובהינתן דופק תקין

מינימלי של  $hr_{mini} = 40bpm$  לפי נוסחה 4.3:

$$n_{mini} = \frac{hr_{mini}[bpm] \cdot 60 \left[ \frac{seconds}{minute} \right]}{f_s[Hz]} \cdot TR[s] = \frac{40[bpm] \cdot 60 \left[ \frac{seconds}{minute} \right]}{900[Hz]} \cdot TR[s] = 2.667TR$$

נוסחה 4.3, חישוב מספר הפסגות המינימלי

ב. מחושב מספר הפסגות המקסימלי  $n_{max}$  כתלות באורך ההקלטה  $TR$  ובהינתן דופק

תקין מקסימלי של  $hr_{maxi} = 150bpm$  לפי נוסחה 4.4:

$$n_{maxi} = \frac{hr_{maxi}[bpm] \cdot 60 \left[ \frac{seconds}{minute} \right]}{f_s[Hz]} \cdot TR[s] = \frac{150[bpm] \cdot 60 \left[ \frac{seconds}{minute} \right]}{900[Hz]} \cdot TR[s] = 10TR$$

נוסחה 4.4, חישוב מספר הפסגות המקסימלי

ג. מוגדר סף ראשוני ושרירותי  $th_0$  למעטפת בערך של 10% מהערך המקסימלי הגלובלי של מעטפת ההקלטה.

ד. מחושב מספר הפסגות הנפרדות העוברות את הסף  $n$ .

ה. במידה ו-  $n < n_{mini}$ , סימן שפספסנו פסגות באמצעות הסף הראשוני, נעדכן את  $th_0$  לפי נוסחה 4.5 ונחזור לסעיף ד'.

$$th_0 \rightarrow 1.01th_0$$

נוסחה 4.5, הגדלת הסף

ו. במידה ו-  $n > n_{maxi}$ , סימן שתפסנו רעשים לא רצויים נוסף על רעשי הלב, נעדכן את  $th_0$  לפי נוסחה 4.6 ונחזור לסעיף ד'.

$$th_0 \rightarrow 0.99th_0$$

נוסחה 4.6, הקטנת הסף

ז. לאחר לכל היותר 10,000 חזרות על סעיפים ד'-ו' נבחר הסף  $th_0$  הנוכחי.

• ראה דוגמה לעדכון חיתוך הסף באיור 4.2.

3. כעת, האלגוריתם מתחשב בתופעה שנקראת *split noise*. לפי תופעה זו, רעשי  $S1$  או  $S2$  עלולים

כל אחד מהם להתפצל לשני רעשים הצמודים זה לזה. הקריטריון להבחנה בין שני רעשים נפרדים

לשתי פסגות מעטפת שנבעו מאותו הרעש הוא מרחק זמני קטן מ- $30ms$ . לכן, מתבצע מעבר על

- מעטפת ההקלטה וכל שתי פסגות הקרובות זו לזו קרבה זמנית הקטנה מ-30ms מסומנות. מתוך כל זוג פסגות האלגוריתם בוחר בפסגה אחת לפי הקריטריון הבא : במידה וה-RMS של השיא השני (המאוחר יותר) גדול בפחות מ-300% של ה-RMS של השיא הראשון (המוקדם יותר), ייבחר השיא הראשון. אחרת, ייבחר השיא השני.
4. בשלב הסרת הפסגות הכפולות מתבצעים הבאים :
- א. חישוב הממוצע וסטיית התקן של המרחקים בין פסגות צמודות.
  - ב. כל עוד סטיית התקן גדולה מערך סף מסוים :
- i. נמצא את המרחק הקטן ביותר בין שתי פסגות
  - ii. נבחר את הפסגה הקטנה ביותר ונוריד אותה מהמעטפת
  - iii. נחזור לסעיף א'

#### שלב שלישי – תיוג שלשות

תיוג הרעשים יתבצע באופן הבא :

1. עבור כל אחת מהפסגות הנותרות לאחר השלב השני :
  - א. נסתכל על המרחק בין הפסגה הנוכחית לפסגה משמאל (זו שקדמה לפסגה הנוכחית) ונסמנו  $d_{left}$
  - ב. נסתכל על המרחק בין הפסגה הנוכחית לפסגה מימין (העוקבת) ונסמנו  $d_{right}$
  - ג. במידה ו- $d_{left} < d_{right}$  נסמן את הפסגה הנוכחית ב-S2. אחרת, נסמנה ב-S1.

שלב זה מסתמך על כך ש- $t_{sys} < t_{dias}$  וראו סקר ספרות.
2. נעבור על תיוגי הפסגות שנמצאו בסעיף הקודם. במידה וקיים קונפליקט בין שתי פסגות שכנות, נסמן כ-unlabeled, כלומר, האלגוריתם מודיע כי לא הצליח להחליט איזה רעש הפסגה מייצגת.

#### 4.1.2 תוצאות

##### ניתוח רוחבי

ביצועי אלגוריתם זה הוערכו על גבי שני מאגרי נתונים. הראשון, *peterjbentley*, הוא מאגר פתוח מתוך אתגר אינטרנטי. מאגר זה מתחלק לשני חלקים : *peterjbentley A* נלקח מהציבור הכללי באמצעות אפליקציית *istethoscope* והחלק השני הוא *peterjbentley B* אשר נלקח מניסויים קליניים באמצעות סטטוסקופ דיגיטלי. כלל מאגר *peterjbentley* מכיל תיוגי מיקומי S1, S2 עבור אנשים בריאים בלבד. המאגר השני הינו מאגר הנתונים הראשון שהתקבל מחברת Sanolla, מאגר זה נלקח מהדגם החדש ביותר של Voqx, הסטטוסקופ החכם של Sanolla, ויקרא מעתה *Sanolla\_v2*. טבלה המסכמת את תוצאות הסגמנטציה על כל אחד מהמאגרים מוצגת בטבלה 4.2 :

Criteria \ Dataset	Peterjbentley A	Peterjbentley B	Sanolla_v2
Unlabeled [%]	20.40	42.61	39.53
Accuracy [%]	95.57	86.98	78.92
Mean Error [ms]	16.51	16.34	22.37



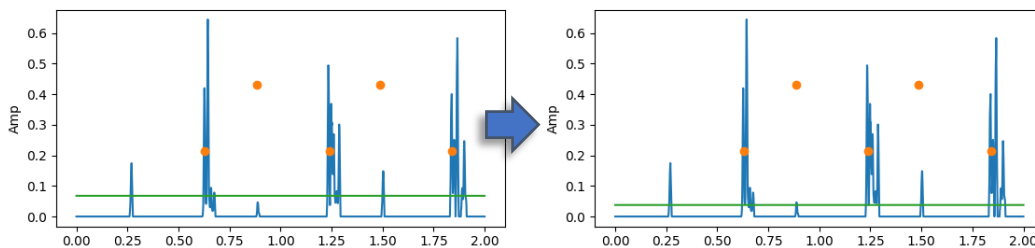
STD Error [ms]	11.95	10.89	14.10
----------------	-------	-------	-------

טבלה 4.2, הערכת ביצועי אלגוריתם הערכת דופק הלב

ניתן לראות כי האלגוריתם אמנם מציג תוצאות טובות עבור אנשים בריאים (מאגר ראשון, *peterjbentley*) אבל התוצאות פוחתות באיכותן באופן משמעותי כאשר ניגשים להקלטות של חברת Sanolla המכילות גם אנשים חולים – אחוזי הדיוק יורדים ואחוזי ה-unlabeled גבוהים גם הם. לאחר הצגת התוצאות לחברת Sanolla הגיעו כותבי פרויקט זה למסקנה כי פתרון לשלב הסגמנטציה בשיטת מעטפת איננו מספק עבור הבעיה הנתונה. זאת, בשל העלמות רעשי  $S2$  בחלק מהתסמונות (וראה סקר ספרות). על כן, הוחלט לממש שיטת פתרון שונה וראה חלק 4.2 שיטת *STFT*.

### דוגמאות ריצה

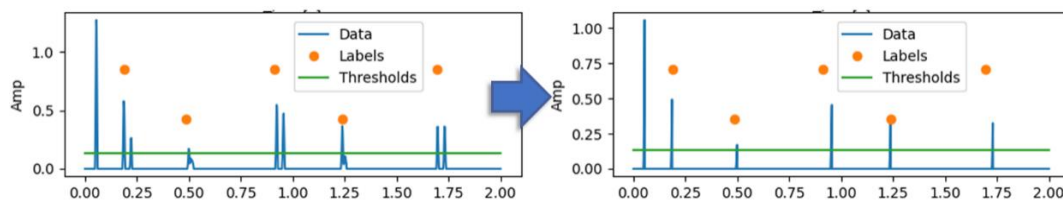
מטרת חיתוך הסף האדפטיבי המתואר ב-4.1.1 היא לזהות פסגות שיתכן ונעלמו בחיתוך הסף הראשוני. דוגמה לשחזור פסגה מוצגת באיור 4.2 :



איור 4.2, שחזור פסגה באמצעות חיתוך סף אדפטיבי

באיור 4.2 מוצגת מעטפת אות נתון. ציר  $x$  הינו ציר הזמן ביחידות של שניות וציר  $y$  מייצגת את משרעת המעטפת ביחידות שרירותיות. בכחול - מעטפת האות, בירוק – הסף הנבחר, בכתום – תגיות מיקומי רעשי  $S1 \setminus S2$ . הגרף משמאל הינו חיתוך הסף עם הסף הראשוני והשרירותי. ניתן לראות כי סף זה מפספס את רעש  $S2$  אשר נמצא ב- $0.9[s]$ . הגרף מימין מייצג את חיתוך הסף הנבחר לאחר עדכון הסף כמתואר ב-4.1.1.

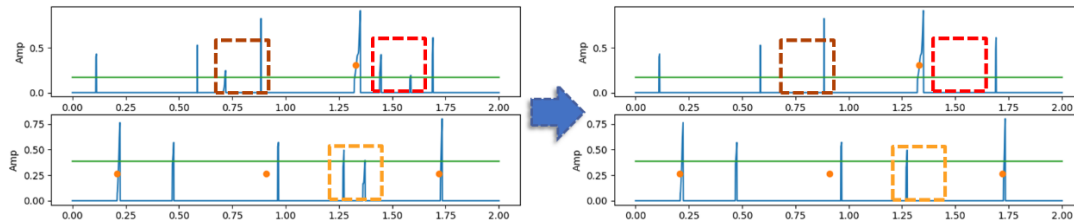
לאחר חיתוך הסף, מתבצעת הורדת *split noises*. דוגמא לכך מוצגת באיור 4.3 :



איור 4.3, הורדת *split noises*

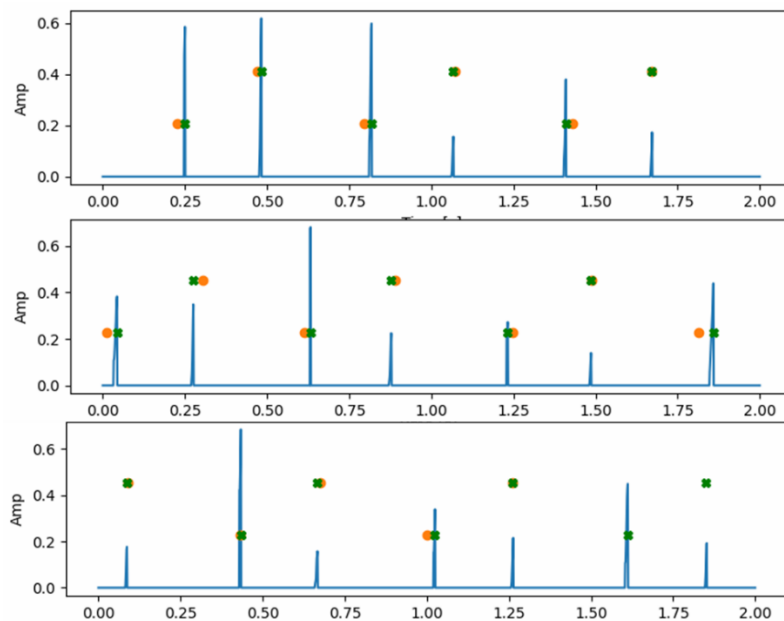
באיור 4.3 מוצגת מעטפת אות נתון. ציר  $x$  הינו ציר הזמן ביחידות של שניות וציר  $y$  מייצגת את משרעת המעטפת ביחידות שרירותיות. בכחול - מעטפת האות, בירוק – הסף הנבחר, בכתום –

תגיות מיקומי רעשי  $S1 \setminus S2$ . הגרף משמאל הינו מעטפת האות לפני הורדת *split noises* והגרף מימין הינו מעטפת האות לאחר הורדת *split noises*. לבסוף, לפני שלב התיג, מתבצעת הורדת פסגות כפולות כמתואר ב-4.1.1. דוגמא לכך מוצגת באיור 4.4:



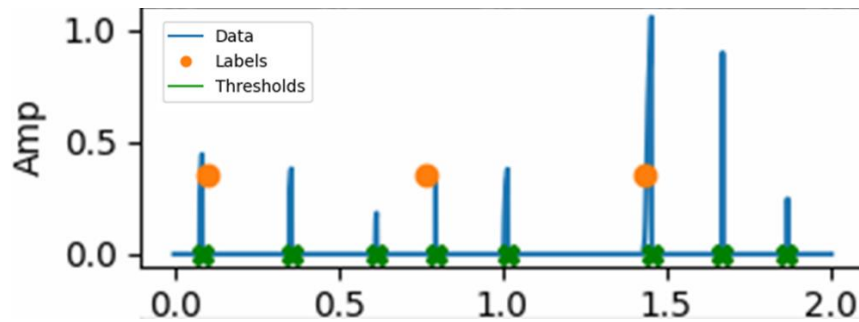
איור 4.4, הורדת פסגות כפולות

באיור 4.4 מוצגת מעטפת אות נתון. ציר  $x$  הינו ציר הזמן ביחידות של שניות וציר  $y$  מייצגת את משרעת המעטפת ביחידות שרירותיות. בכחול - מעטפת האות, בירוק - הסף הנבחר, בכתום – תגיות מיקומי רעשי  $S1 \setminus S2$ . הגרף משמאל הינו מעטפת האות לפני הורדת פסגות כפולות והגרף מימין הינו מעטפת האות לאחר הורדת פסגות כפולות. לנוחות הקורא, הפסגות שנבחרו להורדה מסומנות בריבועים מקווקוים. דוגמא לתיג מוצלח של פסגות מוצגת באיור 4.5:



איור 4.5, תיג פסגות בשיטת מעטפת

באיור 4.5 מוצגת מעטפת אות נתון. ציר  $x$  הינו ציר הזמן ביחידות של שניות וציר  $y$  מייצגת את משרעת המעטפת ביחידות שרירותיות. בכחול - מעטפת האות, בירוק – ניחוי האלגוריתם עבור מיקומי הרעשים, בכתום – תגיות מיקומי רעשי  $S1 \setminus S2$  של *Sanolla*. לעומת איור 4.5, איור 4.6 מציג מקרה בו האלגוריתם לא הצליח לנקות את רעשי ההקלטה ולכן נכשל בסיווג:



איור 4.6, טעות בתיוג פסגות בשיטת מעטפת

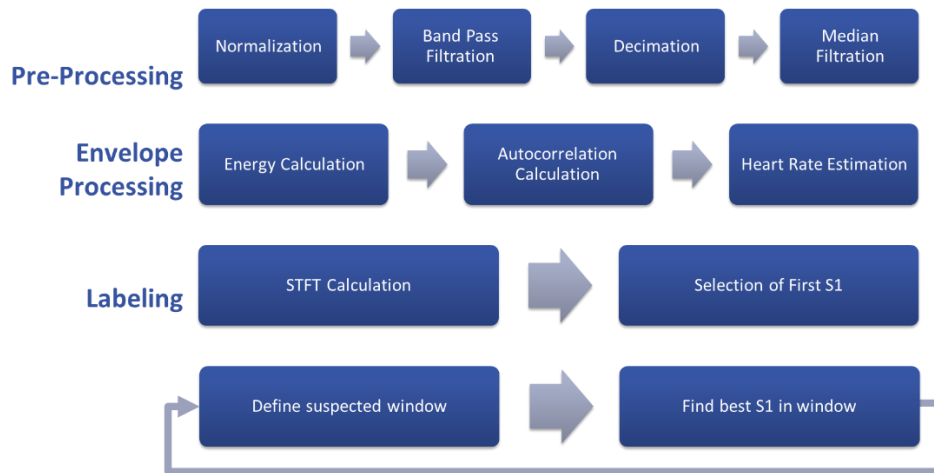
באיור 4.6 מוצגת מעטפת אות נתון. ציר  $x$  הינו ציר הזמן ביחידות של שניות וציר  $y$  מייצגת את משרעת המעטפת ביחידות שרירותיות. בכחול - מעטפת האות, בירוק – ניחושי האלגוריתם עבור מיקומי הרעשים, בכתום – תגיות מיקומי רעשי  $S1 \setminus S2$  של *Sanolla*. ניתן לראות כי רעשי  $S1$  מופיעים במעטפת כשיא הדומה במשרעתו לרעשים נוספים ולא רצויים באות ועל כן הסיווג נכשל. ה-*noise robustness* של האלגוריתם אינו טוב ומהווה סיבה נוספת למעבר לאלגוריתם בשיטת *STFT*.

## 4.2 שיטת STFT

לאחר כישלון שיטת המעטפת בשל הבלעות רעשי  $S2$  בחלק מהאותות שנדרשנו לסווג (וראה חלק 4.1.2, תוצאות שיטת מעטפת), בחרנו לבצע שינוי בהגדרת הבעיה: במקום לחפש את רעשי  $S1$  וגם רעשי  $S2$  בכל הקלטה, בחרנו לחפש את רעשי  $S1$  בלבד. בנוסף, בחרנו לממש שיטת סגמנטציה אחרת, כזו המתחשבת בחתימה התדרית של רעשי הלב ולא רק בעוצמתם ומשכם. שיקולים אלו הובילו אותנו למימוש שיטה מבוססת *STFT* המתוארת בסעיפים הבאים.

### 4.2.1 אלגוריתם

דיאגרמת בלוקים כללית של שלב הסגמנטציה בשיטת *STFT* מוצגת באיור 4.7:



איור 4.7, STFT Segmentation block diagram

### שלב ראשון – עיבוד מקדים

נציין כי הסיגנל שנכנס לאלגוריתם שעליו עבדנו הוא למעשה תוצאת החיסור בין הסיגנל הנקלט על ידי המיקרופון שמכוון כלפי פנים הגוף לבין המיקרופון שמכוון כלפי חוץ. דבר זה נעשה על מנת להקטין את הרעש בסיגנל. בשלב הראשון של הפתרון שמימשנו מתבצע עיבוד מקדים להקלטה לפי השלבים הבאים:

1. ביצוע נורמליזציה לכל הדגימות של המידע המתקבל מהקלטה כך שמתקבל אות בתחום  $[-1, 1]$ .
2. סינון האות באמצעות מסנן 8<sup>th</sup> order Butterworth filter בתחום התדרים  $[25, 400][Hz]$ . תחום זה נבחר מאחר שבסקר הספרות עלה שזהו תחום התדרים שמרכיבים את קולות הלב, ובנוסף מאפשר סינון של רעשי רקע לא רצויים שעלולים להפריע כמו קולות הנשימה המופקים בריאות ומקורם בתדרים גבוהים יותר.
3. דצימציה לתדר של  $f_s = 900[Hz]$  למניעת תופעת התחזות.
4. העברת הסיגנל במסנן חציון עם גרעין בגודל 7 דגימות, כ- $8[ms]$ , לצורך סינון נוסף של רעשים. הפתרון בו מתבצעים חישובי התמרות לצורך מעבר לפתרון במישור התדר.

### שלב שני – מציאת דופק

1. עבור האות שעבר את העיבוד המקדים מחושב האות האנליטי המתואר ע"י נוסחאות 4.1, 4.2.
2. חישוב פונקציית אוטוקורלציה של הערך המוחלט של האות האנליטי לצורך מציאת מחזוריות האות לפי נוסחה 4.7:

$$R_{y_a y_a}[l] = \sum_{n \in \mathbb{Z}} y_a[n] y_a[n - l]$$

נוסחה 4.7, חישוב אוטוקורלציה

כאשר  $R_{xx}[l]$  הינה פונקציית האוטוקורלציה של האות הבדיד והממשי  $y_a = |x_a|$  בזמן  $l$ .

3. לאחר חישוב האוטוקורלציה, נפתח חלון חשוד מימין ל- $R_{y_a y_a}[0]$  אשר קירותיו נקבעים לפי הדופק המהיר ביותר המוגדר כתקין ( $150bpm$ , קיר שמאלי) והדופק האיטי ביותר המוגדר כתקין ( $40bpm$ , קיר ימני). בתוך חלון זה ימצא מיקום המקסימום המקומי של פונקציית האוטוקורלציה. מיקום זה ישמש לחישוב דופק הלב של האדם שהוקלט לפי נוסחה 4.8:

$$hr_e = \text{Estimated Heart Rate} = \frac{60 \cdot f_s}{d} [bpm]$$

נוסחה 4.8, הערכת קצב הלב

כאשר  $f_s = 900[Hz]$  הינו תדר הדגימה של האות המעובד ו- $d$  הינו מספר הדגימות בין האונה הראשית של פונקציית האוטוקורלציה למקסימום הנמצא.

#### שלב שלישי – מציאת ותיוג קולות S1

1. לאחר שיערוך דופק הלב מבוצע חישוב של התמרת STFT על כל הסיגנל עם חלון  $Hann$  בגודל של  $30[ms]$ . לכל נקודה בספקטרוגרמה המתקבלת מחושב הערך המוחלט שלה בריבוע. החישוב מתבצע לפי נוסחה 4.9:

$$X(m, \omega) = \sum_{n=-\infty}^{\infty} x[n]w[n-m]e^{-i\omega n}$$

נוסחה 4.9, חישוב STFT

- כאשר  $X$  התמרת STFT של האות הבדיד  $x$  ו- $w$  הוא חלון  $Hann$ .
2. מתוך ההתמרה שחושבה בשלב הקודם מתבצע חיפוש של דגימות שבסבירות גבוהה משתייכות לקול מסוג S1. החיפוש מתבצע באמצעות העברה של "מסכה" על פני הספקטרוגרמה וחישוב "ניקוד" של כל דגימה בזמן ע"י כפל איבר-איבר בין ה"מסכה" לבין החלק בספקטרוגרמה אותו היא מכסה.
- ה"מסכה" הייתה בעלת ממדים המתאימים למשך הזמן המקסימלי הידוע עבור קול מסוג S1 לפי הספרות,  $15[ms]$ , ובנוסף יתאימו לחתימה התדרית של רעש  $S1$ ,  $10 - 200[Hz]$ .
3. לאחר העברת המסכה על פני הספקטרוגרמה שהתקבלה, נבחרת הדגימה בעלת הניקוד הגבוה ביותר להיות דגימה המתאימה ל-S1 בסבירות הגבוהה ביותר. דגימה זו משמשת כעוגן, נקודת התחלה לצורך מציאת שאר מיקומי רעשי S1.
4. כעת האלגוריתם נכנס לשלב איטרטיבי:

- א. מוגדר חלון חשוד הממוקם במיקום  $hr_e$  דגימות מימין לעוגן הנוכחי בגודל המתאים לאורכו של רעש S1 אופייני בנוסף ל- $10[bpm]$  המתאים לשגיאה מוערכת בהערכת הדופק של  $\pm 5[bpm]$ .
- ב. בחלון זה מתבצע חיפוש אחר הדגימה המתאימה ביותר לרעש S1 באופן זהה לחיפוש בשלב 2.
- ג. הדגימה המתאימה ביותר לרעש S1 מוגדרת כעוגן החדש ונשמר מיקומו.
- ד. במידה ולא הגיע סוף האות, חזרה לשלב 4.א'

5. שלב 5 זהה לשלב 4 למעט שינוי קל – החיפוש מתבצע מהעוגן הראשוני (אשר נמצא בשלב 3).  
לכיוון שמאל (לתחילת האות).
6. העוגנים שנמצאו במהלך שלבים 3-5 נשמרים כמיקומי S1 בהקלטה.

## 4.2.2 תוצאות

### ניתוח רוחבי

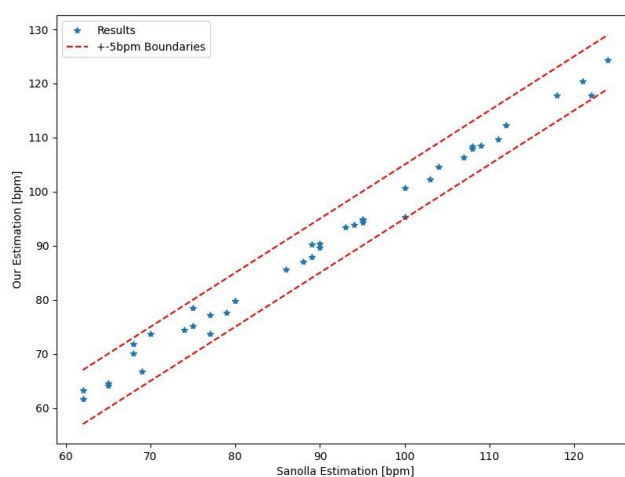
את מאגר הנתונים הראשון שהתקבל ע"י חברת Sanolla, *Sanolla\_v2*, ניתן לחלק לשני חלקים אשר שניהם מורכבים מהקלטות דופק של חולים שנלקחו מאותו דגם של מכשיר, כאשר באחד החלקים (מעתה *trainHR*) הומלץ לנו ע"י החברה להשתמש לצורך הערכה של שיטת שיערוך הדופק שלנו. שני החלקים מכילים תגיות הן של קצב הלב המוערך ע"י Sanolla והן של מיקומי S1 בכל הקלטה.

הערכה מדויקת של קצב הלב מוגדרת ככזו אשר נמצאת בטווח של  $\pm 5bpm$  מהערכת חברת Sanolla. תוצאותינו מוצגות בטבלה 4.3:

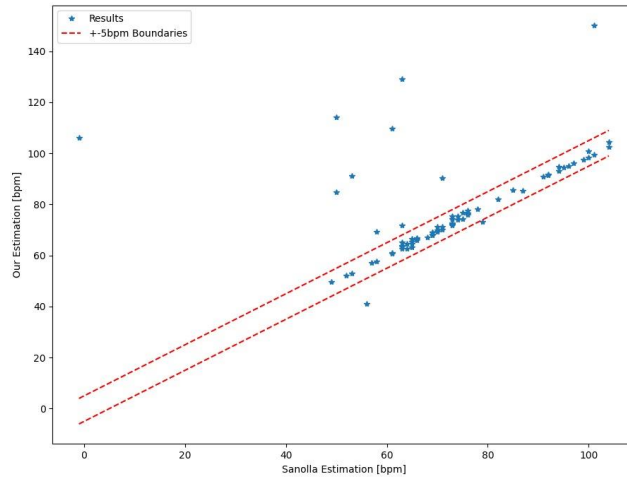
Criteria \ Dataset	trainHR	Rest
Accuracy	100%	86%
Mean Error	1.1bpm	6.8bpm
STD Error	1.3bpm	18.1bpm

טבלה 4.3, הערכת ביצועי אלגוריתם הערכת דופק הלב

אילוסטרציה של תוצאות אלו מוצגות באיורים 4.8, 4.9:



איור 4.8, הערכת ביצועים להערכת קצב לב, *trainHR*



איור 4.9, הערכת ביצועים להערכת קצב לב, Rest of sanolla\_v2 dataset

באיורים 4.8, 4.9 ציר  $x$  מייצג את הערכת Sanolla וציר  $y$  את הערכת האלגוריתם שלנו, שניהם ביחידות של  $bpm$ . כל כוכב כחול מייצג הקלטה נפרדת וקווי ה- $tolerance$  באדום מייצגים טווח של  $\pm 5bpm$ . ניחוש נכון לרעש  $S1$  נחשב כזה אשר נמצא במרחק של עד  $70[ms]$  מתגית של Sanolla. הערכת ביצועי הסגמנטציה על מאגר הנתונים Sanolla\_v2 מוצגות בטבלה 4.4:

Criteria \ Dataset	trainHR	Rest
Accuracy Per Label	96%	74%
Accuracy Per Recording	95%	72%
Mean Error	17[ms]	69[ms]

טבלה 4.4, הערכת ביצועי אלגוריתם הסגמנטציה על מאגר Sanolla\_v2

ניתן לראות שעל החלק במאגר הנתונים של החברה המכונה trainHR התקבלו תוצאות טובות עם אחוזי דיוק גבוהים, כאשר המרחק בין הסיווגים שהתקבלו לאחר הרצת האלגוריתם שלנו לזה של Sanolla לבין אלה שקיבלנו מהחברה קרוב מאוד ועומד על  $17[ms]$  בממוצע. לעומת זאת, ניתן לראות מהטבלה שכאשר עוברים לחלקים האחרים, בהם ההקלטות נלקחו בבתי חולים חלה ירידה בדיוק האלגוריתם ויש ירידה משמעותית בתוצאות. לא הצלחנו להגיע למסקנה חד משמעית לגבי הסיבות לירידה בדיוק האלגוריתם, אך אנו סבורים שאחת הסיבות שעשויה לתרום לכך היא איכות ההקלטות שנלקחו, והעובדה שהחלק הראשוני של האלגוריתם שבו מתבצע סינון הרעשים לא מצליח לנטרל את הרעשים בצורה מספיק טובה, היא גורם שתורם לירידה באחוזי הדיוק של האלגוריתם. בוצע פילוח של הסיבות לטעויות הסגמנטציה. תוצאות פילוח זה מוצגות בטבלה 4.5.

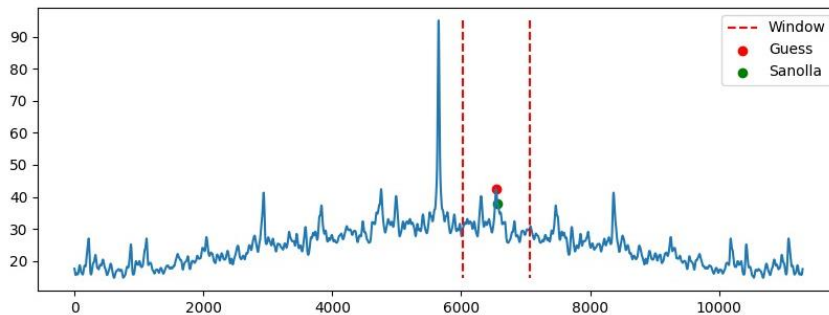
סיבת הטעות	אחוז התגיות הלא נכונות
ניחוש ראשון לא נכון	50%
הערכת קצב לב שגויה	29%

8%	S1 רחב או ניחוש מוקדם
3%	קצב לב לא רגולרי
10%	לא ידוע

טבלה 4.5, פילוח טעויות סגמנטציה

## דוגמאות ריצה

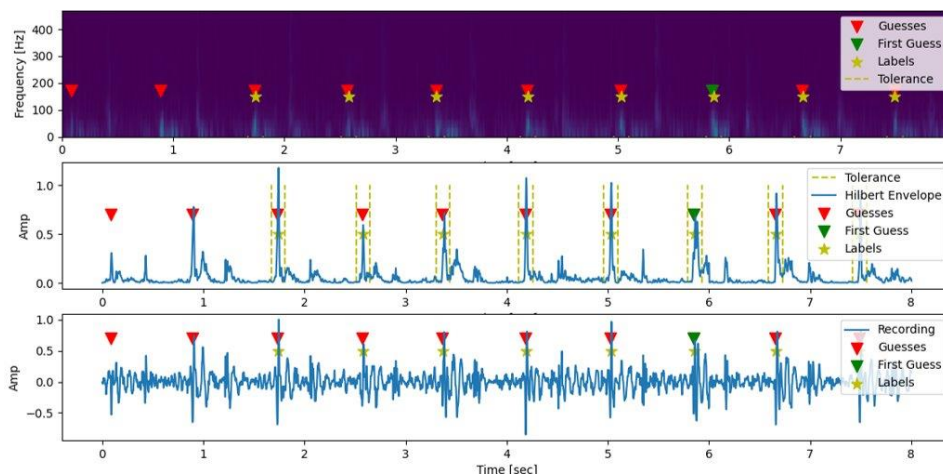
דוגמא להערכת קצב הלב ניתן לראות באיור 4.10:



איור 4.10, דוגמא להערכת קצב הלב

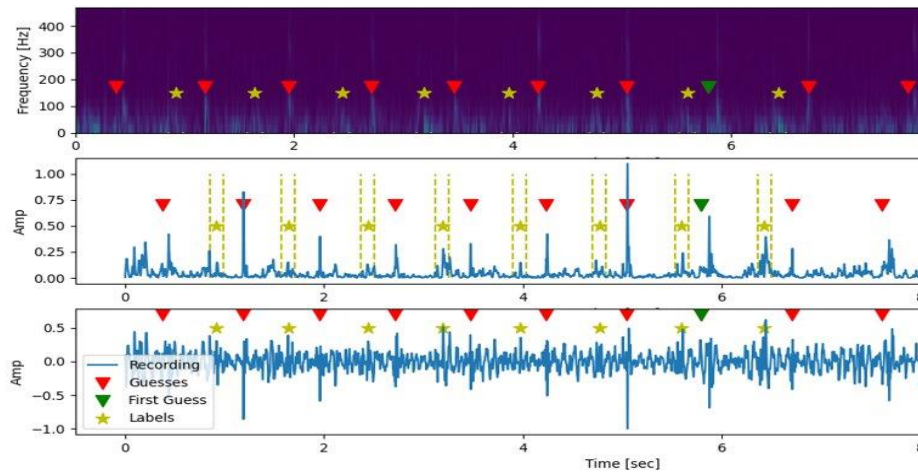
באיור 4.10 מוצגת האוטוקורלציה של אחת ההקלטות במאגר, ציר  $x$  הינו ביחידות של דגימות וציר  $y$  מייצג את משרעת האוטוקורלציה ביחידות שרירותיות. בשלב הראשון מחושב חלון חשוד (מקווקו באדום) אשר קירו השמאלי מייצג את הדופק המהיר ביותר המוגדר כתקין ( $150bpm$ ) וקירו הימני מייצג את הדופק האיטי ביותר המוגדר כתקין ( $40bpm$ ). בתוך חלון זה נמצא הדגימה בה האוטוקורלציה מגיע למקסימום (עיגול אדום) ומתוך דגימה זו מוערך דופק הלב לפי נוסחה 4.8. לצורך השוואה מוצגת הערכת *Sanolla* עבור אות זה (עיגול ירוק).

דוגמאות מייצגות עבור שתי הקלטות מוצגות באיורים 4.11 ו-4.12:



איור 4.11, דוגמה לסגמנטציה טובה





איור 4.12, דוגמה לסגמנטציה לא טובה

בכל אחד מהאיורים ניתן לראות שלושה גרפים. למעלה, התמרת  $STFT$  של האות, ציר  $x$  הוא ציר הזמן ביחידות של  $[s]$ , ציר  $y$  הוא ציר התדר ביחידות של  $[Hz]$ , צבע כחול בתמונה מתאים למשרעת גבוהה וצבע סגול למשרעת נמוכה. באמצע, מעטפת האנרגיה של האות כפי שחושבה בשלב השני של 4.2.1, ציר ה- $x$  הוא ציר הזמן ביחידות של  $[s]$  וציר  $y$  הוא משרעת המעטפת  $[A \cdot U]$ . למטה, האות שנכנס לאלגוריתם הסגמנטציה לאחר עיבוד מקדים, ציר  $x$  הוא ציר הזמן ביחידות של  $[s]$  וציר  $y$  הוא משרעת המעטפת  $[A \cdot U]$ . על גבי כל אחד משלושת הגרפים מוצגים: הסיווג עבור רעש  $S1$  הראשון – משולש ירוק, שאר הסיווגים – משולשים אדומים, תגיות שסופקו מחברת Sanolla – כוכבים צהובים, ו- $tolerance$   $70[ms]$  מסביב לתגיות של Sanolla – קווים מקווקים צהובים.

באיור 4.11 ניתן לראות כי האלגוריתם הצליח להשיג תוצאות קרובות לתגיות של Sanolla – הניחוש הראשון נפל על אחת התגיות, דופק הלב הוערך כהלכה ושאר רעשי  $S1$  נראים באופן מספיק ברור בספקטרוגרמה כך שהאלגוריתם הצליח לזהותם. לעומת זאת, באיור 4.12 ניתן לראות כי האלגוריתם לא הצליח להשיג תוצאות קרובות לתגיות של Sanolla – אמנם דופק הלב הוערך כהלכה אבל הניחוש הראשון נופל על רעש דומיננטי בהקלטה שאיננו רעש  $S1$  (סביר כי מדובר ברעש  $S2$ ) ולכן כל שאר הניחושים רחוקים גם הם בהסחה קבועה מתגיות Sanolla.

## 5. בעיית הסיווג

הקלט של בעיית הסיווג הוא מחזור בודד אשר מתקבל כפלט של בעיית הסגמנטציה. פלט הסיווג הוא מספר בינארי המציין אם המחזור שייך להקלטה של אדם בעל קולות לב נורמליים או לא-נורמליים.

עבור בעיית הסיווג נבחנו 3 פתרונות שונים. ראשית, נבחן תיאור של מחזור ע"י מספר מצומצם של פיצורים בשילוב עם מסווג  $SVM$  (5.1). שנית, נבחן תיאור של מחזור ע"י מספר רב של מקדמי  $MFC$  ובניית עץ החלטה (5.2). לבסוף, נבחן תיאור של מחזור ע"י ספקטרוגרמת  $MFC$  בשילוב עם רשת נוירונים בסגנון  $ResNet$  (5.3).

כל אחת מהשיטות לעיל מומשו והוערכו על מספר מאגרי נתונים. שני מאגרי הנתונים הראשונים הם  $sanolla\_v2$  ו- $peterjbentley$  אשר תוארו בהרחבה בפתיחת פרק 4. נוסף על שני מאגרים אלו, הוערכו ביצועי המסווגים על מאגר  $sanolla\_v1$  אשר מכיל הקלטות שנרכשו ע"י דגם ישן של מכשיר Voqx מבית חברת Sanolla. מאגר  $sanolla\_v1$  גדול יותר משני המאגרים הקודמים ולכן שימש כמאגר נוח לאימון ובחינת המסווגים. טבלה 5.1 משווה בין מאגרי הנתונים השונים ששימשו את הפרויקט לשלב הסיווג.

מאגר	<i>peterjbentley</i>	<i>sanolla_v1</i>	<i>sanolla_v2</i>
מקור	אתגר אינטרנטי	חברת Sanolla	חברת Sanolla
מכשיר הקלטה	<i>iStethoscope</i>	<i>Voqx</i>	<i>Voqx</i>
מס' הקלטות	312	723	318
מס' הקלטות סימפטומטיות	112	231	11

טבלה 5.1, מאגרי נתונים לשלב הסיווג

### 5.1 פיצורים תיאוריים ו- $SVM$

דיאגרמת בלוקים כללית של מסווג זה מופיעה באיור 5.1.



איור 5.1, דיאגרמת בלוקים של מסווג  $SVM$

#### 5.1.1 חילוץ פיצורים

מתוך מחזור בודד חולצו 10 מספרים המתארים את המחזור לפי הפירוט הבא:

**Zero Crossing Rate** .א

קצב חציית האפס של האות  $x_t$  לפי נוסחה 5.1 :

$$zcr = \frac{1}{T-1} \sum_{t=1}^{T-1} I\{x_t x_{t-1} < 0\}$$

נוסחה 5.1, חישוב  $zcr$

כאשר  $I\{\cdot\}$  הוא אינדיקטור ו- $T$  מספר הדגימות באות.

#### ב. **MFCC Mean, STD**

שני מספרים המייצגים את הממוצע וסטיית התקן של מקדמי ה- $MFC$  של האות. להרחבה בנושא מקדמי  $MFC$  ראה חלק 5.2.1, חילוף פיצורים.

#### ג. **Spectral Centroid**

מיקום מרכז המסה של הספקטרום של האות. חישוב זה מתבצע בדומה לחישוב מרכז מסה מכני – את המיקום מחליף ציר התדר בהתמרת  $DFT$  של האות והמסה מוחלפת ע"י משרעת ההתמרה בכל נקודה על ציר התדר. תיאור פורמלי מוצג בנוסחה 5.2.

$$sc = \frac{\sum_{n=0}^{N-1} f[n]x[n]}{\sum_{n=0}^{N-1} x[n]}$$

נוסחה 5.2, *Spectral Centroid*

כאשר  $f[n]$  ו- $x[n]$  הם משרעת ההתמרה והתדר בדגימה  $n$  בהתאמה.

#### ד. **Spectral Rolloff**

התדר עבורו 85% מהאנרגיה הספקטרלית נמצאת מתחתיו.

#### ה. **Spectral Flux**

שטף ספקטרילי הינו מדד למהירות השינוי של ספקטרום העוצמה של האות. מדד זה מחושב ע"י ביצוע השוואה (בד"כ  $L2 - norm$ ) בין ספקטרום האות ב- $frame$  אחד לזה העוקב.

#### ו. **Mean Frequency**

מספר מרוכב (שני פיצורים) המתארים את התדר הממוצע במחזור הקלט.

#### ז. **Energy Entropy**

חישוב של אנטרופיית מעטפת האות האנליטי מבוסס התמרת הילברט אשר תואר בחלק 4.1.1.

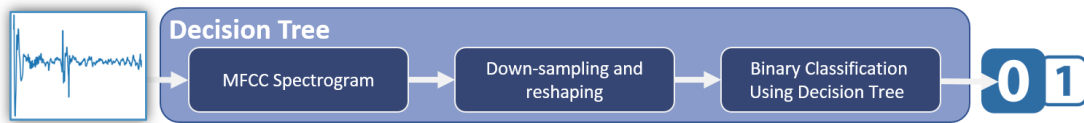
#### ח. **Pad Number**

כל המחזוריים במאגר הנתונים מרופדים באפסים לאורך קבוע ארוך מספיק בשביל להתאים למחזור הגדול ביותר האפשרי. מספר האפסים הנדרש לריפוד קשור לקצב הלב שהוערך עבור ההקלטה אבל הינו ייעודי וייחודי לכל מחזור.

עשרת הפיצ'רים אשר תוארו מעלה שימשו את המערכת לתיאור כל מחזור נתון. כל אחד מהמאגרים הנתונים המתוארים בפתיחת פרק זה עברו את שלב הסגמנטציה בשיטת  $STFT$ , המחזורים אשר התקבלו חולקו ל-Train\Test split 80\20 ושימשו לאימון והערכת ביצועים של מסווג מסוג  $SVM$ .

## 5.2 מקדמי MFC ועצי החלטה

דיאגרמת בלוקים כללית של מסווג זה מופיעה באיור 5.2.



איור 5.2, דיאגרמת בלוקים של מסווג  $DT$

### 5.2.1 חילוץ פיצ'רים

מתוך מחזור בודד חולצו פיצ'רים מסוג  $MFC - Mel - Frequency Cepstrum$ . פיצ'רים אלו נמצאים בשימוש במערכות  $speech recognition$  וגם בסיווג אותות לב. יתרונם היחסי על מפות פיצ'רים מבוססות זמן-תדר אחרות כמו  $STFT$  או  $Wavelets$  היא בהתאמת הספקטרום לאופן בו אוזן האדם תופסת קול, ספקטרום זה נקרא  $Mel Spectrum$  וההמרה בינו לבין ציר התדר הסטנדרטי ניתן לחישוב מתוך נוסחה 5.3.

$$Mel(f) = 2595 \log_{10} \left( 1 + \frac{f}{700} \right)$$

נוסחה 5.3, חישוב  $Mel Spectrum$

חישוב הפיצ'רים עבור מחזור נתון מתבצע לפי השלבים הבאים:

#### א. $Pre - emphasis$

סינון  $HPF$  פרמטרי של אות הכניסה לפי נוסחה 5.4:

$$y[n] = x[n] - \alpha x[n - 1]$$

נוסחה 5.4,  $Pre - emphasis$

כאשר  $x[n]$  אות הכניסה,  $\alpha$  פרמטר בתחום  $[0,1]$  (בפרויקט זה  $\alpha = 0.97$ ) ו- $y[n]$  אות המוצא.

## ב. Framing

חלוקה של כל מחזור ל- $Frames$  בגודל של כ- $30[ms]$  באמצעות חלונות  $Hann$  לפי נוסחה 5.5 :

$$W(n) = (1 - \beta) - \beta \cos\left(\frac{2\pi n}{N-1}\right), \quad 0 \leq n \leq N-1$$

נוסחה 5.5,  $Hann$  Framing של מחזור

כאשר  $\beta = 0.46$  ו- $N$  מספר הדגימות ב- $Frame$ .

## ג. DFT Power Spectrum

חישוב עוצמת התמרת ה- $DFT$  של  $P(k)$  של כל  $Frame$  לפי נוסחה 5.6 :

$$P(k) = \frac{1}{N} |X(k)|^2 = \frac{1}{N} \left| \sum_{n=0}^{N-1} x[n] e^{-\frac{2\pi j n k}{N}} \right|^2, \quad 0 \leq n, k \leq N-1$$

נוסחה 5.6, ספקטרום העוצמה של התמרת  $DFT$

כאשר  $X(k)$  הינה התמרת  $DFT$  של  $x[n]$  ו- $N$  מספר הדגימות ב- $Frame$ .

## ד. Log Spectrum

חישוב ספקטרום האנרגיה הלוגריתמי  $S(m)$  יתבצע לפי נוסחה 5.7 :

$$S(m) = \ln \left[ \sum_{k=0}^{N-1} P(k) H_m(k) \right], \quad 0 \leq m \leq M$$

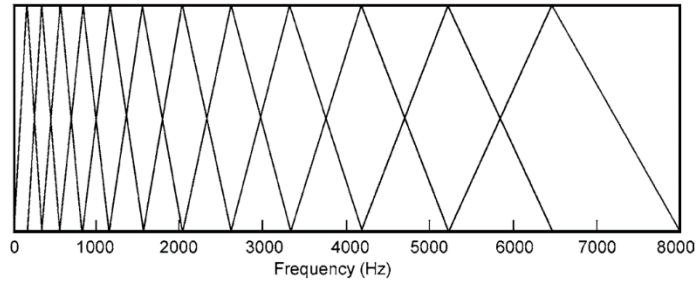
נוסחה 5.7, ספקטרום אנרגיה לוגריתמי

כאשר  $M$  פרמטר הקובע את מספר מקדמי מל, בפרויקט זה נבחר  $M = 13$ ,  $P(k)$  הוא ספקטרום העוצמה של התמרת ה- $DFT$  אשר מחושב לפי נוסחה 5.6, ו- $H_m(k)$  הינם מסנני מל – מסננים משולשים בעלי תמך ותדר בסיס משתנה אשר ניתנים לחישוב לפי נוסחה 5.8 :

$$H_m(k) = \begin{cases} 0, & k < f(m-1) \\ \frac{k - f(m-1)}{f(m) - f(m-1)}, & f(m-1) \leq k \leq f(m) \\ \frac{f(m+1) - k}{f(m+1) - f(m)}, & f(m) \leq k \leq f(m+1) \\ 0, & k > f(m+1) \end{cases}$$

נוסחה 5.8, חישוב מסנני מל

כאשר  $f(l)$  הוא תדר מל ה- $l$  אשר נקבע באמצעות נוסחה 5.3. לשם המחשה, איור 5.3 מציג מספר מסנני מל מסדרים נמוכים.



איור 5.3, מסנני מל

## ה. DCT

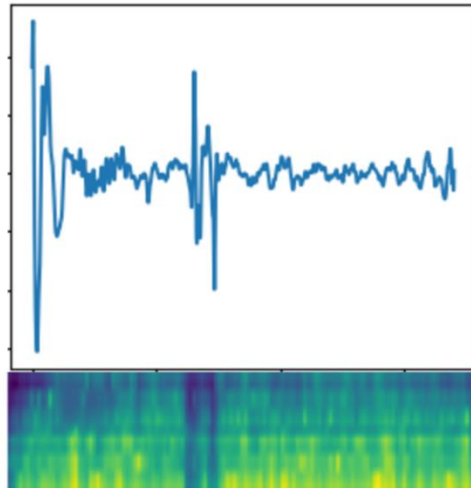
ספקטרום האנרגיה הלוגריתמי עובר התמרת DCT לפי נוסחה 5.9 :

$$C(n) = \sum_{m=0}^{M-1} S(m) \cos\left(\frac{\pi n \left(m - \frac{1}{2}\right)}{M}\right), \quad n = 1, 2 \dots L$$

נוסחה 5.9, DCT on Lof Spectrum

כאשר  $M = L = 13$ .

לאחר ביצוע שלבים א' – ה', כל מחזור מתואר ע"י מטריצה בגודל  $13 \times 64$  כאשר 13 הוא מספר מקדמי מל בכל  $Frame$  ו-64 הוא מספר ה- $Frames$ . דוגמא למטריצה כזו, מעתה ספקטרוגרמת MFC, מוצגת באיור 5.4.



איור 5.4, ספקטרוגרמת MFC

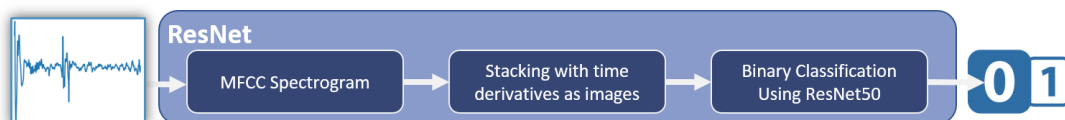
באיור 5.4 שתי תמונות נפרדות. למעלה, מחזור בודד, מוצא של מערכת הסגמנטציה. למטה, ספקטרוגרמת MFC המתאימה למחזור הנתון, ציר x הוא ציר הזמן וציר y הוא ציר תדר המל, צהוב מסמל עוצמה גבוהה וכחול עוצמה נמוכה. חילוץ הפיצ'רים למסווג זה מסתיים בשינוי צורת הספקטרוגרמה לווקטור באורך 832 פיצ'רים.

## 5.2.2 סיווג

832 הפיצ'רים אשר תוארו מעלה שימשו את המערכת לתיאור כל מחזור נתון. כל אחד מהמאגרים הנתונים המתוארים בפתיחת פרק זה עברו את שלב הסגמנטציה בשיטת *STFT*, המחזורים אשר התקבלו חולקו ל-Train/Test split 80\20 ושימשו לאימון והערכת ביצועים של מסווג מסוג *Decision Tree*.

## 5.3 ספקטרוגרמת MFC ורשת נוירונים

דיאגרמת בלוקים כללית של מסווג זה מופיעה באיור 5.5.



איור 5.5, דיאגרמת בלוקים של מסווג ResNet

## 5.3.1 חילוץ פיצ'רים

מתוך מחזור בודד חולצו פיצורים מסוג *Mel – Frequency Cepstrum - MFC*. פיצורים אלו נמצאים בשימוש במערכות *speech recognition* וגם בסיווג אותות לב. יתרונם היחסי על מפות פיצורים מבוססות זמן-תדר אחרות כמו *STFT* או *Wavelets* היא בהתאמת הספקטרום לאופן בו אוזן האדם קולטת קול, ספקטרום זה נקרא *Mel Spectrum* וההמרה בינו לבין ציר התדר הסטנדרטי ניתן לחישוב מתוך נוסחה 5.3. חילוף הפיצורים למסווג זה מתבצע בשני שלבים: ראשית, חולצה ספקטרוגרמת *MFC* של כל מחזור לפי המתואר ב-5.2.1, שלבים א'-ה'. שנית, הספקטרוגרמה המתקבלת עוברת גזירה בציר הזמן פעמיים לפי נוסחה 5.10, שתי התוצאות הנקראות *Dynamic Features*, הינן גם הן מטריצות בגודל  $64 \times 13$  בדומה לספקטרוגרמת ה-*MFC*.

$$\begin{cases} D(n) = \frac{1}{\sqrt{\sum_{i=-k}^k i^2}} \sum_{i=-k}^k iC(n+i) \\ D_2(n) = \frac{1}{\sqrt{2 \sum_{i=-k}^k i^2}} \sum_{i=-k}^k iD(n+i) \end{cases}$$

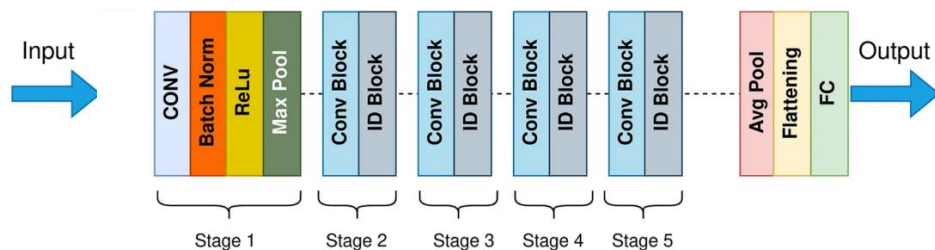
נוסחה 5.10, חישוב *Dynamic Features*

התוצאה הסופית הינה "תמונה" סינטטית בעלת שלושה ערוצים ( $C, D, D_2$ ) ובגודל של  $64 \times 13$ . תמונה זו היא הקלט של הרשת המתוארת ב-5.3.2.

### 5.3.2 סיווג

#### ארכיטקטורה

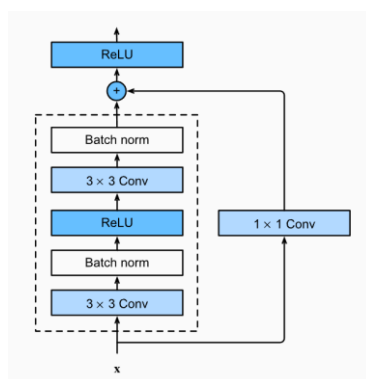
תמונת הפיצורים בגודל  $3 \times 64 \times 13$  שימשה את המערכת לתיאור כל מחזור נתון. כל אחד מהמאגרים הנתונים המתוארים בפתיחת פרק זה עברו את שלב הסגמנטציה בשיטת *STFT*, המחזורים אשר התקבלו חולקו ל-Train/Test split 80\20 ושימשו לאימון והערכת ביצועים של רשת נוירונים בעלת ארכיטקטורת *ResNet50*. מבנה הרשת מוצג באיור 5.6.



איור 5.6, ארכיטקטורת רשת *ResNet50*



כאשר 5 – 2 Stages המורכבים מ-Conv Block + ID Block הם דרגות Residual blocks סטנדרטיות המורכבות משכבות קונבולוציה 3X3 ושרשרת התוצאה לקונבולוציה 1X1. מבנה ה-Residual Block מוצג באיור 5.7.



איור 5.7, Residual Block

### בחירת היפר-פרמטרים

ההיפר-פרמטרים של הרשת הינם גודל ה-Batch, קצב הלימוד ומספר ה-Epochs. קבוצת האימון חולקה חלוקה פנימית ל-80/20 כאשר 20% הם ה-Validation Set. התבצע אימון מלא של הרשת עבור כלל הפרמוטציות מבין הערכים הבאים:

(Batch Size = 64, 32, 16; Learning Rate =  $10^{-3}$ ,  $10^{-4}$ ,  $10^{-5}$ ; Epochs = 10, 15, 20)

ההיפר-פרמטרים אשר הביאו לביצועים הטובים ביותר הם:

(Batch Size = 64, Learning Rate =  $10^{-4}$ , Epochs = 15)

### הערכת ביצועים

עבור ההיפר-פרמטרים שנבחרו לעיל, הרשת אומנה באופן מלא וביצועיה הוערכו על גבי קבוצת המבחן.

## 5.4 תוצאות

### 5.4.1 חיזוי בהינתן מחזור בודד

לאור הכמות הקטנה של הדוגמאות הסימפטומטיות במאגר sanolla\_v2 (וראה טבלה 5.1), הרשת המתוארת ב-5.3 לא אומנה על מאגר זה. למעט חריגה זו, כלל המאגרים שימשו לאימון והערכת הביצועים של כלל המסווגים. טבלה 5.2 מסכמת את תוצאות החיזוי בהינתן מחזור יחיד. הערכים בטבלה מופיעים בפורמט (F1 Score) Accuracy %.

מאגר נתונים	SVM	Decision Tree	ResNet
peterjbentley	70 (63)%	67 (66)%	75 (72)%

64 (63)%	54 (52)%	59 (54)%	<i>sanolla_v1</i>
NA	90 (1)%	97 (0)%	<i>sanolla_v2</i>
70 (42)%	61 (42)%	68 (16)%	<i>combined dataset</i>

טבלה 5.2, סיכום תוצאות הסיווג בהינתן מחזור בודד

אחוזי ה-*F1 Score* הנמוכים עבור מאגר *sanolla\_v2* והמאגר המשולב נובעים ממספר הדוגמאות הסימפטומטיות הקטן של מאגר *sanolla\_v2* בהתאם להערה לעיל. התוצאות הטובות ביותר הושגו עבור מסווג ה-*ResNet* והגרועות ביותר עבור מסווג ה-*Decision Tree*. בנוסף, ניתן לראות שהמאגר שסווג באופן הטוב ביותר הוא המאגר *peterjbentley*.

#### 5.4.2 חיזוי בהינתן הקלטה

נוסף על תוצאות החיזוי בהינתן מחזור יחיד, נבדקו התוצאות גם בהינתן הקלטה שלמה. כלומר, המסווג אומן על קבוצה של מחזורים וחה סיווג לכל מחזור בודד ולאחר מכן בוצע *majority vote* בין כלל תוצאות החיזוי המשתייכות לאותה ההקלטה. תוצאות פעולה זו מוצגות בטבלה 5.3. הערכים בטבלה מופיעים בפורמט Accuracy (F1 Score) %.

<i>ResNet</i>	<i>Decision Tree</i>	<i>SVM</i>	מאגר נתונים
80 (77)%	71 (69)%	76 (67)%	<i>peterjbentley</i>
69 (65)%	65 (59)%	69 (63)%	<i>sanolla_v1</i>
NA	92 (1)%	97 (0)%	<i>sanolla_v2</i>
74 (48)%	68 (47)%	73 (24)%	<i>combined dataset</i>

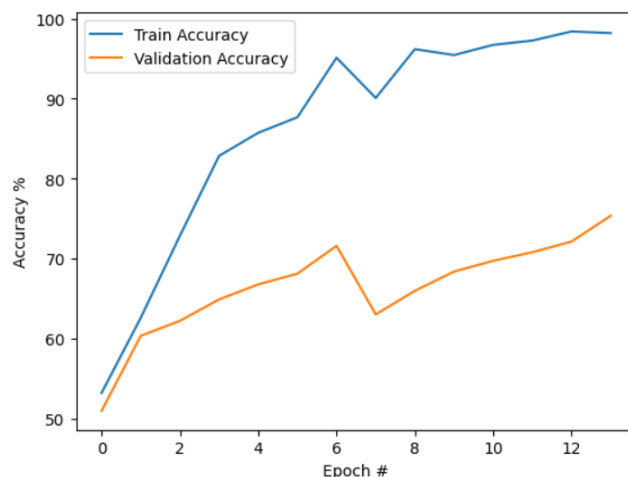
טבלה 5.3, סיכום תוצאות הסיווג בהינתן הקלטה

אחוזי ה-*F1 Score* הנמוכים עבור מאגר *sanolla\_v2* והמאגר המשולב נובעים ממספר הדוגמאות הסימפטומטיות הקטן של מאגר *sanolla\_v2* בהתאם להערה לעיל. התוצאות הטובות ביותר הושגו עבור מסווג ה-*ResNet* והגרועות ביותר עבור מסווג ה-*Decision Tree*. בנוסף, ניתן לראות שהמאגר שסווג באופן הטוב ביותר הוא המאגר *peterjbentley*. תוצאות אלו מראות שיפור של 3-10% באחוזי הדיוק בין כלל המאגרים והמסווגים בהשוואה לתוצאות בהינתן מחזור בודד ב-5.4.1. התוצאות הטובות ביותר הינן 80% דיוק עבור מסווג ה-*ResNet* ומאגר *peterjbentley*.

#### 5.4.3 הרחבה – אימון מודל ה-*ResNet*

מודל ה-*ResNet* הינו מודל מורכב מאוד באופן יחסי לשני המסווגים הבסיסיים יותר אותם בחנו. למודל ישנם מעל 23 מיליון פרמטרים ניתנים לאימון. בהתאם לכך, תוצאות החיזוי על קבוצת האימון הינן גבוהות והרשת מבצעת *Overfitting* על קבוצה זו. דוגמה לתופעה זו ניתן

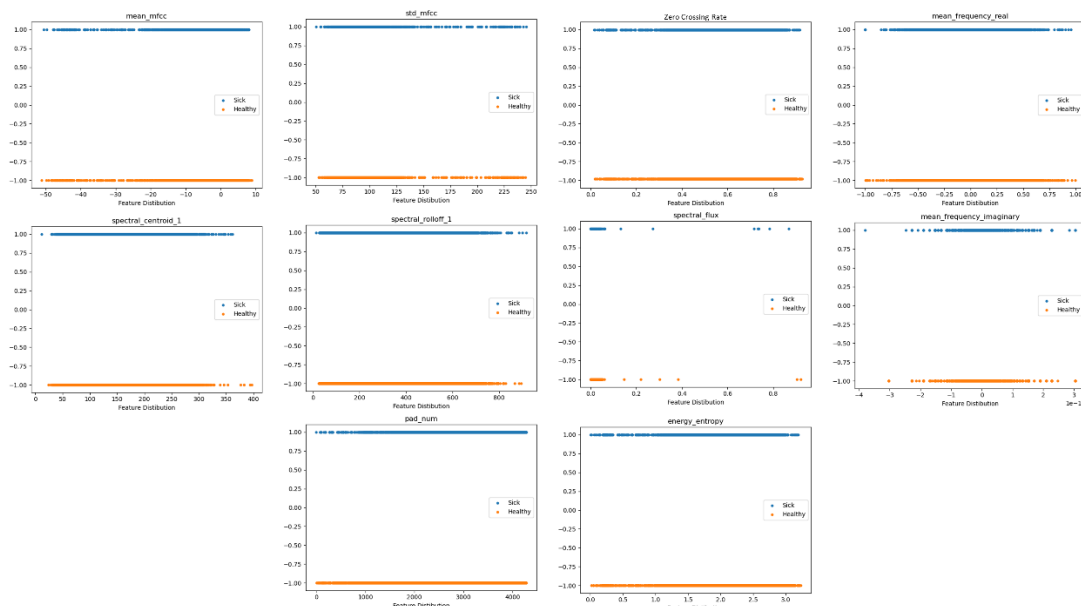
לראות באיור 5.8. הדוגמה נלקחה מתוך אימון הרשת על המאגר *sanolla\_v1* בהינתן ההיפר-פרמטרים ב-5.3.2.



איור 5.8, ResNet Overfitting

#### 5.4.4 הרחבה - הפרדה ע"י פיצורים תיאוריים

נבחנה יכולתם של פיצורים בודדים להפריד בין הקלטה של קולות נורמליים ללא-נורמליים. תוצאותינו מוצגות באיור 5.9.



איור 5.9, הפרדה ע"י פיצורים תיאוריים

באיור 5.9 10 גרפים, אחד לכל פיצור. שם הפיצור מופיע בכותרת כל גרף. בכחול - מחזורים חולים, בכתום – מחזורים בריאים. ציר y מייצג את ההפרדה בין חולים (1) לבריאים (-1). ציר x מייצג את פיזור הערכים עבור הפיצור הנתון.

מתוך איור 5.9 ניתן לראות כי אף פיצ'ר בודד לא מפריד בצורה טובה בין מחזורים בריאים למחזורים חולים – התחומים חופפים כמעט לחלוטין בין שתי המחלקות.

## 6. מסקנות

1. עבודה עם מאגרי מידע רפואיים – במהלך הפרויקט עבדנו עם מאגרי מידע רפואיים אשר הציגו בעיות "עולם אמיתי" אשר נדרשנו להתמודד איתן. בתוך כך, ההקלטות נפתחו בדרך כלל ברעש חזק שנבע מהנחת המיקרופון ותחילת ההקלטה, במקרים רבים רעשי סביבה נוספים כמו דיבורים, ונקיפת מתכות הופיעו גם באמצע ההקלטה. במאגרי המידע של *sanolla* בהם עבדנו עם מידע גולמי נעזרנו בשני המיקרופונים הקיימים בסטוסקופ החכם על מנת לנסות ולהתגבר על רעשי הסביבה. ניקוי רעשים משני מיקרופונים לא היה במיקוד הפרויקט ולכן לא הושקעה בכך עבודה רבה. להערכת כותבי פרויקט זה, זו הסיבה העיקרית להבדלים בין התוצאות במאגר *sanolla* ומאגר *peterjbentley*.
2. סגמנטציה של אותות לב – סגמנטציה באמצעות שיטות מעטפת איננה שיטה טובה מספיק על מנת להתמודד עם הקלטות של קולות לב לא-נורמליים ובנוסף גישה זו מתקשה בטיפול ברעשים. גישה הסתברותית המתארת את האות באמצעות *STFT* הראתה תוצאות מצוינות עבור אנשים בריאים ותוצאות טובות-בינוניות עבור אנשים חולים או הקלטות רועשות. להערכתנו ניתן לשפר את שיטת הסגמנטציה באופן משמעותי באמצעות מספר שינויים ועל כך יורחב בפרק 8. יצוין כי הערכת הביצועים של שיטת הסגמנטציה איננה שלמה – מספר הדוגמאות המתויגות של קולות לב לא-נורמליים הינו קטן מכדי להגיע למסקנות חותכות ונדרשת עבודת מחקר נוספת.
3. סיווג בינארי של אותות לב – סיווג באמצעות מודל של רשת נוירונים בעלת ארכיטקטורת *ResNet* המקבלת מפת פיצ'רים מבוססת *MFCC* הראה את התוצאות הטובות ביותר מבין כלל האפשרויות שנבחנו. מודל זה הינו מורכב ואקספרסיבי במיוחד ועל כן הרשת מבצעת *Overfit* לקבוצת האימון ונדרשת עבודת המשך על מנת לטפל בכך.

## 7. סיכום

בפרויקט זה התמודדנו בהנחיית מעבדת SIPL ובשיתוף עם חברת Sanolla עם בעיית סיווג אותות לב. פתרון בעיה זו יאפשר לחסוך זמן של רופאים מומחים הנדרשים לבצע אוסקלציה ידנית, ינגיש יכולת לאבחון ראשוני של מחלות לב לכלל הציבור וימנע טעויות אנוש.

במסגרת פרויקט זה נבחנה גישה שונה של סיווג אותות לב מהגישה המופיעה בספרות – אימון מסווג בהינתן מחזור בודד. לצורך כך, מומשו הן מערכת סגמנטציה והן מערכת סיווג נפרדות ועצמאיות. במהלך הפרויקט נבחנו מספר פתרונות שונים לכל אחד משלבי הבעיה כאשר הפתרון הטוב ביותר הגיע ל-80% דיוק בסיווג. תוצאה זו איננה עומדת ב-*state of the art* לפי הספרות המקצועית (90% ומעלה) אך אנו מאמינים שהגישה הנבחרת בעלת הפוטנציאל להגיע לכך וראה פרק 8, עבודה עתידית.

במהלך כתיבת הפרויקט פיתחו הכותבים את יכולתם הן בעיבוד אותות קלאסי והן במסווגים ולמידה עמוקה. צברנו ניסיון בעבודה עם מאגרי נתונים רועשים, אלגוריתמים לחלוקה של אותות וסוגים שונים של פיצ'רים לתיאור אותות קול.

## 8. עבודה עתידית

1. הערכה מחודשת של התוצאות על מאגר *Sanolla* נקי :  
נדרשת עבודה מעמיקה בנושא ניקוי רעשים משני המיקרופונים בשימוש הסטטוסקופ  
החכם של חברת *Sanolla* והערכה מחודשת של תוצאות האלגוריתם הקיים לאחר מכן.
2. ניחוש ראשון בבעיית הסגמנטציה :  
כפי שניתן לראות בטבלה 4.5, כ-50% מטעויות בתגיות הסגמנטציה נבעו מניחוש ראשון  
לא נכון, בדרך כלל ניחוש של  $S2$  כ- $S1$ . בעיה זו ניתנת לפתרון ע"י עבודה נוספת על  
הפרדה בין  $S1$  ל- $S2$  או באמצעות חתימתם התדרית או באמצעות רוחבם הזמני של  
הרעשים או חלקי המחזור שהם פותחים. להרחבה על הבדלים אלו ראה טבלה 3.1.
3. שיפור מודל ה-*ResNet* :  
מודל ה-*ResNet* הראה את התוצאות הטובות ביותר ולאחר טיפול בשתי הנקודות לעיל  
אנו ממליצים להכניס במודל זה מספר שיפורים :  
א. שימוש באורך המחזור – נכון לכתיבת שורות אלו, ההבדלים באורך המחזור אינם  
באים לידי ביטוי כפיצ'ר בכניסה לרשת, כלל המחזורים מרופדים באפסים לאורך  
זזה ומתוכם מחולצות מפות הפיצ'רים כמתואר ב-5.3.1.  
ב. קצב למידה אדפטיבי – שימוש בקצב למידה המתעדכן במהלך האימון על מנת  
למנוע חלק מה-*Overfitting* שמבצעת הרשת על קבוצת האימון.

1. S. Sun, Z. Jiang, H. Wang, and Y. Fang, "Automatic moment segmentation and peak detection analysis of heart sound pattern via short-time modified Hilbert transform," *Comput. Methods Programs Biomed.* vol. 114, no. 3, pp. 219-230, 2014.
2. H. Liang, S. Lukkarinen, I. artimo, "Hear Sound Segmentation Algorithm Based on Heart Sound Envelopogram", *Computers in Cardiology*, Vol 24, 1997.
3. D. Kumar, P. Carvalho, M. Antunes, J. Henriques, L. Eugenio, R. Schmidt, J. Habetha, "Detection of S1 and S2 Heart Sounds by High Frequency Signatures", *EMBS Annual international conference*, 2006.
4. Amit Krishna Dwivedi, Syed Anas Imtiaz, Esther Rodriguez-Villegas, "Algorithms for automatic analysis and classification of hear sounds – a systematic review", *IEEE access* 2018.
5. Feng Li, Zheng Zhang, Lingling Wang, Wei Liu, "Heart sound classification based on improved mel-frequency spectral coefficients and deep residual learning" *Front. Physiol.*, 22 December 2022