

**דו"ח סיכום פרויקט: א'**

זיהוי הפרעות קול

Voice interference detection

מבצעים:

**דרור אריה Dror Arie**

**אור באייר Or Bayer**

מנחה: **[שם המנחה] [Supervisor Name]**

**סמסטר רישום: חורף תשע"ח**

**תאריך הגשה: ‏יולי, 2019**

**בשיתוף עם: [שם חברה] [לוגו חברה]**

****

ניתן לרשום כאן הקדשה, למשל:

אנו מקדישים דו"ח מסכם זה למשפחותינו האהובות.

**תודות**

דף אופציונלי שבו אתם יכולים להודות לגורמים שונים, למשל:

אנו מודים מקרב לב לחברת XXX שהציעה את הרעיון לפרויקט ולאיש הקשר בחברה YYYY שתמך בנו במסירות בכל מהלכו.

אנו מודים גם ליאיר משה שכתב את תבנית זו לדו"ח מסכם ובכך עזר לנו לכתוב דו"ח מסכם בהצלחה ☺

**תוכן עניינים**

[1.](#_heading=h.gjdgxs) הוראות לכתיבת דו"ח 1

[1.1.](#_heading=h.3znysh7) מבנה כללי 1

[1.2.](#_heading=h.2et92p0) פורמט 2

[2.](#_heading=h.3dy6vkm) פרק לדוגמה 3

[3.](#_heading=h.3rdcrjn) כתיבת מאמר אקדמי 5

[4.](#_heading=h.35nkun2) שימוש במונחים בעברית 7

[5.](#_heading=h.1ksv4uv) סיכום 11

[רשימת מקורות 12](#_heading=h.44sinio)

**רשימת איורים**

[איור 1. סמל המעבדה לעיבוד אותות ותמונות. 3](#_heading=h.1t3h5sf)

[איור 2. גרף לדוגמה.](about:blank) 3

**רשימת טבלאות**

[טבלה 1. המרה ממעלות צלסיוס לפרנהייט. 4](#_heading=h.4d34og8)

**תקציר**



**Abstract**

# הוראות לכתיבת דו"ח

מסמך זה מהווה תבנית לכתיבת דו"ח מסכם לפרויקט במעבדה לעיבוד אותות ותמונות. ניתן ואף רצוי לכתוב את הדו"ח תוך כדי שינוי של מסמך זה, שכבר כתוב בפורמט המומלץ.

## מבנה כללי

הדו"ח מתחיל בדף שער סטנדרטי. את דף השער יש למלא במקומות המתאימים - שם הפרויקט ושמות הסטודנטים והמנחה בעברית ובאנגלית, שם החברה ולוגו שלה (במידה וקיימת לפרויקט חברה מלווה). בנוסף, יש לבחור את הערכים המתאימים בקופסאות הבחירה - סוג פרויקט, סמסטר ושנה, תאריך הגשה, מספר ארכיון. מספר הארכיון מורכב ממספר הפרויקט במערכת LabAdmin (ארבע ספרות), מספר הסמסטר (1 – חורף, 2 – אביב, 3 – קיץ) ושנה (שתי ספרות).

לאחר דף השער ניתן לרשום הקדשה ודף תודות ואז מגיעים לתוכן העניינים, רשימת איורים ורשימת טבלאות (במידה וקיימות). רשימות אלה ניתן ליצור באופן אוטומטי באמצעות הלשונית References בקבוצות Table of Contents ו-Captions בהתאמה. לשם יצירת תוכן עניינים תקין, יש להקפיד שכל הכותרות בדו"ח יהיו תוך שימוש בסגנון Heading. לשם יצירת רשימת איורים ורשימת טבלאות, יש להקפיד ליצור כתוביות לכל האיורים והטבלאות. עוד על כך בהמשך. טיפ: יצירת תוכן עניינים בפורמט תקין בעברית עלולה להסתבך מעט. כדי לפתור בעיות מסוג זה, עליכם לשנות את הגדרות הטאבים בתוכן – רצוי לקבוע טאב אחד מיושר לימין עבור תחילת הכותרות וטאב אחד מיושר לשמאל עבור מספרי העמודים.

התקציר מתאר בקצרה את הפרויקט, הן בעברית והן באנגלית. אורכו חצי עמוד.

לב הדו"ח מורכב מפרקים ותתי-פרקים ממוספרים, לפי שיקול דעתכם. לרב, מתחילים בפרקי מבוא וסקר ספרות, ממשיכים לתיאור הפתרון ומסיימים בתוצאות ומסקנות. כל פרויקט לגופו. הפרק האחרון הוא פרק סיכום קצר שחוזר על התקציר עם מעט יותר פרטים ותוך התייחסות למה שנאמר בגוף הדו"ח.

לסיום, רשימת מקורות, שגם אותה ניתן ליצור אוטומטית. עוד על כך בהמשך.

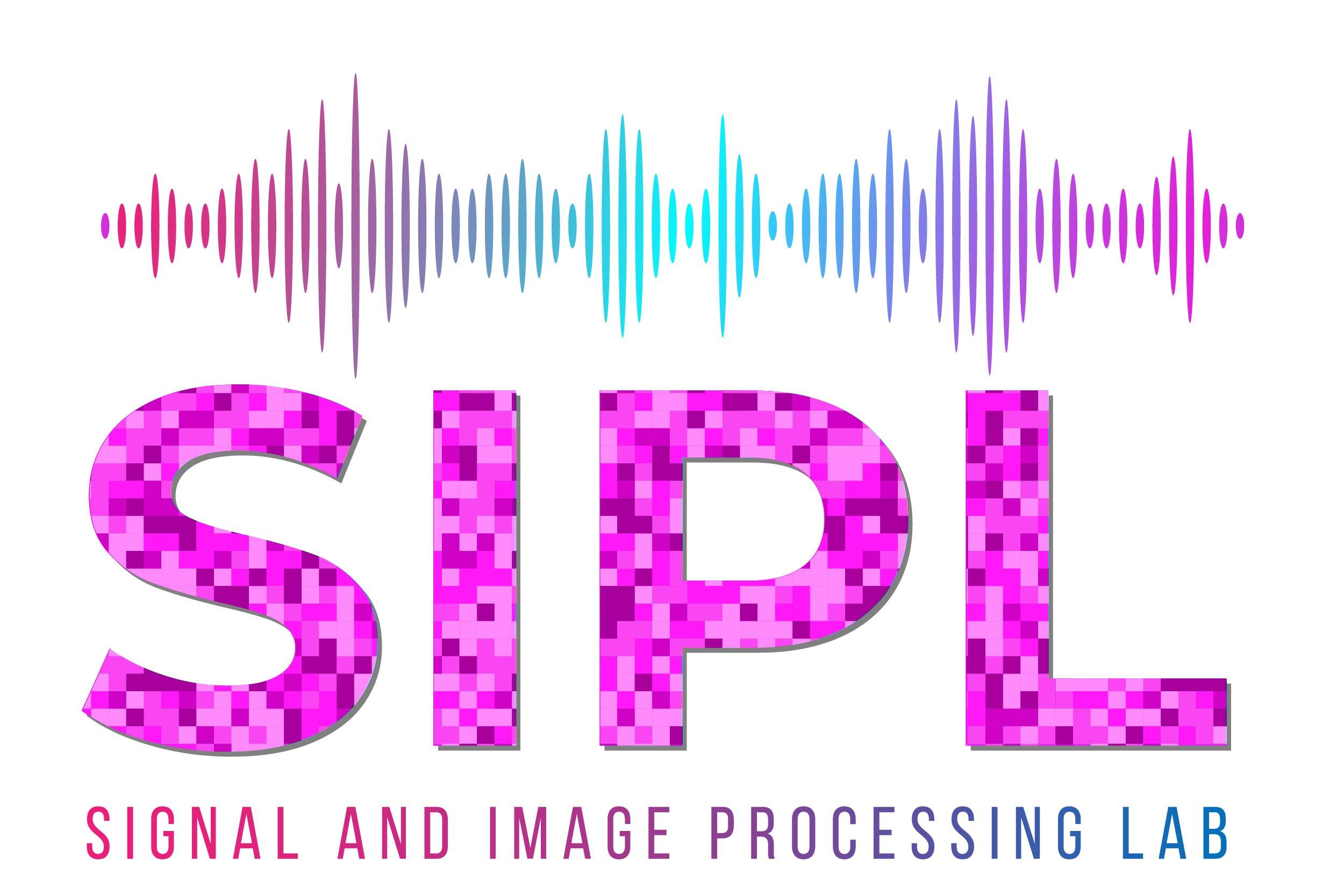
## פורמט

מסמך זה כתוב בפורמט המומלץ כך שהדרך הפשוטה ביותר היא להשתמש בו כבסיס לכתיבת הדו"ח. הקפידו על יישור הטקסט הן לימין והן לשמאל וכתבו בצורה מסודרת, ברורה ורשמית יחסית. שימו לב לא למחוק בטעות את מספור העמודים בחלק התחתון של כל עמוד.

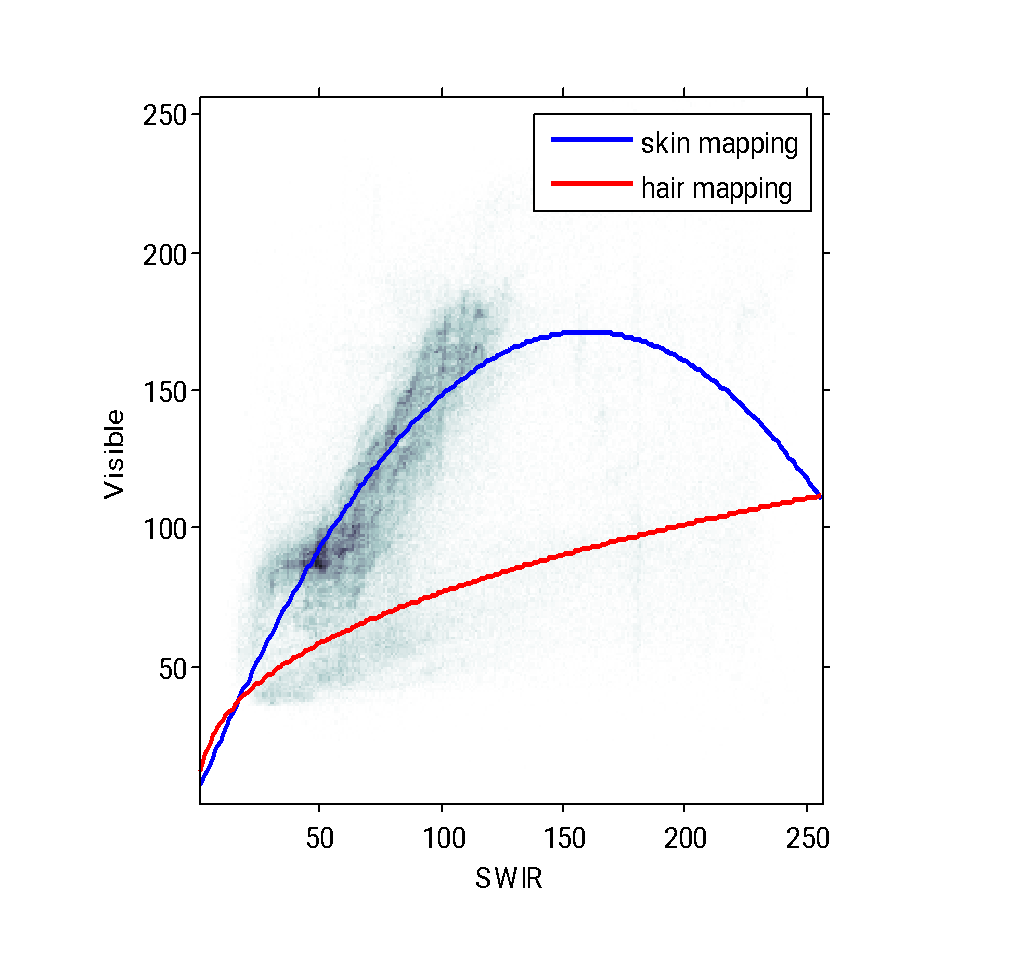
# פרק לדוגמה

מטרת פרק זה היא להוות דוגמה לכתיבת טקסט בגוף הדו"ח, תוך שימוש נכון באיורים, טבלאות והפנייה למקורות.

כל גרף, איור, תמונה וכד' צריכים להיות מלווים בכתובית ויש להתייחס אליהם במפורש בטקסט. כדי לגרום לכתוביות להיות בפורמט אחיד וכדי ליצור מהן בהמשך רשימת איורים, יש להוסיף אותן באמצעות Insert Caption בקבוצה Captions בלשונית References ולהתייחס אליהן בטקסט בעזרת האפשרות Cross-reference בקבוצה זו. הנה דוגמה להתייחסות לאיור: כפי שרואים באיור 1, סמל המעבדה לעיבוד אותות ותמונות מכיל את המילה SIPL. שימו לי כי המילים איור 1 במשפט הקודם ובמשפט זה מהוות קישור וכך מספר האיור ישתנה אוטומטית בטקסט עם הוספת איורים נוספים.



*איור 1. סמל המעבדה לעיבוד אותות ותמונות.*

כאשר מוסיפים לדו"ח גרפים שיוצרו ב-MATALB, כמו באיור 2, יש להקפיד שהטקסט בגרף גדול מספיק ושהתמונה יוצאה מ-MATLAB באופן וקטורי (פורמט EMF או EPS). באופן זה, איכות הטקסט והקווים בתמונה תישמר גם בהגדלה שלה.

מה שנאמר לגבי איורים נכון גם לגבי טבלאות אלא שכאן נהוג לרשום את הכתובית מעל הטבלה. לדוגמה, טבלה 1.

*טבלה 1. המרה ממעלות צלסיוס לפרנהייט.*

| Fahrenheit | Celsius |
| --- | --- |
| 68 | 20 |
| 77 | 25 |
| 86 | 30 |

נוסחאות ניתן לרשום באמצעות Equation בלשונית Insert. למשל, היא התמרת DFT בדידה של סדרה :

במקרה של נוסחאות מורכבות במיוחד, ניתן להשתמש ב-MathType אך לרב אין זה נדרש.

בסוף הדו"ח יש לכלול רשימת מקורות בפורמט של IEEE, כפי שמודגם במסמך זה. יש להקפיד על רשימת כל הפרטים הרלוונטיים עבור כל מקור (שמות המחברים, שם המאמר או הספר, מקום פרסום, שנה, ...). יש להתייחס לכל המקורות בטקסט והם צריכים להופיע ברשימת המקורות בסדר שבו נעשית ההתייחסות אליהם בטקסט. Word יכול לעזור לנו גם בזה. הוא מאפשר לנהל רשימת מקורות ולהפנות אליהם בקישורים בטקסט בפורמט המתאים בעזרת הפעולות בלשונית References תחת הקבוצה Citations & Bibliography. לדוגמה, הנה התייחסות לספר שמסביר כיצד לכתוב מאמר אקדמי [1], הנה אזכור של מאמר כנס ומאמר ירחון [2] [3], והנה אזכור לאתר אינטרנט [4]. אם ניהול רשימת המקורות הופך מורכב, ניתן להשתמש בכלי חיצוני שמתווסף ל-Word כגון הכלי החינמי Mendeley.

# כתיבת מאמר אקדמי

בחלק קטן מהפרויקטים במעבדה, כאשר יש לכך הצדקה אקדמית ונכונות מצד הסטודנטים, כותבים הסטודנטים ביחד עם המנחה מאמר לכנס בינלאומי. פרק זה מציין מספר כללים מנחים לכתיבת מאמר אקדמי והוא רלוונטי רק לסטודנטים שכותבים מאמר כזה.

באופן היסטורי, נעשה שימוש נרחב ב-LaTeX לכתיבת מאמרים אקדמיים וגם כיום הוא כלי נפוץ. עם השנים התפתחו כלי עריכה נוחים יחסית עבורו, כמו Overleaf. עם זאת, ניתן כיום לכתוב מאמרים אקדמיים בנוחות גם בעזרת Word בהתאם להוראות במסמך זה לכתיבת דו"ח מסכם. את המאמר כותבים כמובן באנגלית תוך שימוש בתבנית שמסופקת על ידי מארגני הכנס. מדובר לרב על 4 עד 6 עמודים בשתי עמודות בגופן קטן יחסית בפורמט IEEE:

<https://www.ieee.org/conferences/publishing/templates.html>

הכללים הבאים שצוינו קודם במסמך זה נוגעים גם למאמר אקדמי:

* כתיבת תקציר, פרקי "לב העבודה", פרק תוצאות ופרק מסקנות
* הוספת איורים וטבלאות ממוספרים עם כתוביות מתאימות
* רשימת מקורות בפורמט IEEE

במאמר אקדמי נהוג להצמיד את האיורים לשוליים העליונים של העמוד. אפשרי, אך פחות נהוג, להצמיד לשוליים התחתונים של העמוד. בכל מקרה אין להכניס את האיורים לא בצמוד לשוליים העליונים או התחתונים. כלל זה אינו נוגע לטבלאות, שאותן ניתן לשלב בטקסט. במידה ונדרשים איור או טבלה גדולים, ניתן לפרוס אותם לרוחב שתי העמודות. ברירת המחדל של Word היא לעגן כל איור לשורת טקסט ולהזיז אותו כאשר שורת הטקסט זזה. כדי למנוע מאיור או מכתובית לזוז, יש לבצע את הפעולות הבאות (לכל אחד מהם בנפרד): לאחר בחירת האיור או הכתובית, בלשונית Format בקבוצה Arrange ללחוץ על Wrap Text ושם לבחור More Layout Options. בחלון שנפתח, בלשונית Text Wrapping יש לבחור Top and Bottom. בלשונית Position יש לבחור, גם עבור Horizontal וגם עבור Vertical, יישור שהוא Absolute position בהתייחס ל-Page. בחירה זו תקבע את האיור במקום קבוע בעמוד אך הוא עדיין עלול לזוז למקום זהה בעמוד אחר, בהתאם לטקסט שאליו הוא מעוגן. כדי לקבוע לאיזו שורת טקסט מעוגן כל איור, יש לבחור בתפריט File את האפשרות Options ושם תחת Display להפעיל את האפשרות Object anchors. עכשיו, בכל פעם שתלחצו על איור או כתובית, תראו צלמית קטנה של עוגן ( ) שמציינת לאיזו שורת טקסט מעוגן האיור. ניתן לגרור את העוגן כרצונכם. בכל מקרה, כדאי לוודא שכל איור והכתובית שלו מעוגנים שניהם לאותה שורת טקסט כדי שיזוזו יחד.

בנוגע לנוסחאות, ההמלצה להשתמש בכלי המובנה של Word עומדת בעינה. במאמרים אקדמיים נהוג למספר כל נוסחה ולהפנות אל המספור במקום המתאים בטקסט. היכולת למספר נוסחאות אינה נתמכת באופן ישיר ב-Word אך ניתן לעשות זאת על ידי תחימת הנוסחה בטבלה בלתי-נראית, כפי שרואים ב-(1).

| (1) |  |
| --- | --- |

הסבר כיצד לעשות זאת ניתן למצוא בקישור הבא:

<https://word.tips.net/T000273_Numbering_Equations.html>

# שימוש במונחים בעברית

את הדו"ח המסכם אתם נדרשים לכתוב בעברית (חוץ מאשר במקרה חריגים בהם יש לקבל אישור מיוחד מצוות המעבדה). בכתיבה מקצועית בעברית אין מנוס משימוש במונחים מקצועיים רבים באנגלית. עם זאת, קיימים לא מעט מונחים מקצועיים בתחום שלנו שיש להם תחליף טוב בעברית. במקרים אלה אתם מתבקשים להשתמש במונח העברי. להלן רשימה קצרה בסדר אלפביתי של מונחים מקצועיים באנגלית והתחליף שלהם בעברית.

| המונח המקביל בעברית | המונח באנגלית |
| --- | --- |
| דיוק | accuracy |
| מסתגל | adaptive |
| ניתוח | analysis |
| תוצא שווא | artifact |
| ניחות | attenuation |
| מסנן מעביר פס | band-pass filter |
| רוחב פס | bandwidth |
| סיבית, קצב סיביות | bit, bitrate |
| תרשים מלבנים | block diagram |
| תיבה תוחמת | bounding box |
| צוואר בקבוק | bottleneck |
| כיול | calibration |
| ערוץ | channel |
| סיווג | classification |
| אשכול | cluster |
| תצורה | configuration |
| ניגודיות | contrast |
| קו מתאר | contour |
| נתונים | data |
| מאגר נתונים, מאגר מידע, מסד נתונים | database, dataset |
| מפענח | decoder |
| למידה עמוקה | deep learning |
| ברירת מחדל | default |
| תכן | design |
| השהיה | delay |
| גילוי, גלאי | detection, detector |
| עיוות | distortion |
| תחום | domain |
| שפה, גילוי שפות | edge, edge detection |
| איבר | element |
| לשכן, שיכון | embed, embedding |
| מקודד | encoder |
| אומדן | estimation |
| משוב | feedback |
| מאפיין, מיצוי מאפיינים, מפת מאפיינים | feature, feature extraction, feature map |
| מסנן, סינון | filter, filtering |
| מסגרת | frame |
| קצב מסגרות | frame rate |
| פס תדרים | frequency band |
| כללי | global |
| סריג | grid |
| נתוני אמת | ground truth |
| מסנן מעביר גבוהים | high-pass filter |
| עוצמת הארה | illumination |
| דימות | imaging |
| מימוש | implementation |
| הסק | inference |
| אתחול | initialization |
| קלט | input |
| גרעין | kernel |
| תגית, תיוג | label, labelling |
| זמן המתנה | latency |
| שכבה | layer |
| מקומי | local |
| איכון | localization |
| בהיקות | luminance |
| פונקציית הפסד | loss function |
| מסנן מעביר נמוכים | low-pass filter |
| למידת מכונה | machine learning |
| יריעה | manifold |
| חציון | median |
| מזעור | minimization |
| אפנון | modulation |
| רשת עצבית | neural network |
| עצם | object |
| חיזוי | prediction |
| התאמת יתר | overfitting |
| חריג | outlier |
| פלט | output |
| שיא | peak |
| עיבוד מקדים | preprocessing |
| אקראי | random |
| קצב | rate |
| גולמי | raw |
| זמן אמת | real-time |
| זיהוי | recognition |
| יתירות | redundancy |
| למידה על ידי חיזוקים | reinforcement learning |
| דגימה מחדש | resampling |
| היענות | response |
| חסין, חסינות | robust, robustness |
| רוויה | saturation |
| חיישן | sensor |
| פריד | separable |
| אות | signal |
| יחס אות לרעש | signal to noise ratio (SNR) |
| בו-זמני | simultaneous |
| למידה מפוקחת | supervised learning |
| תבנית | template |
| סדרת בוחן | testing set |
| עקיבה | tracking |
| סדרת אימון | training set |
| סף, ביצוע פעולת סף | threshold, thersholding |
| התמרה | transformation |
| כוונון | tuning |
| למידה לא מפוקחת | unsupervised learning |
| סדרת אימות | validation set |
| שונות | variance |
| משקל | weight |

# סיכום

כאן יש להוסיף סיכום קצר של הפרויקט (מה הייתה המטרה וכיצד השגנו אותה) ומסקנות תוך התייחסות למה שנאמר בגוף הדו"ח.

שיהיה בהצלחה ☺

# רשימת מקורות

| [1] | B. Gustavii, How to Write and Illustrate a Scientific Paper, 2nd ed., Cambridge, UK: Cambridge University Press, 2008. |
| --- | --- |
| [2] | E. Bullkich, I. Ilan, Y. Moshe, Y. Hel-Or and H. Hel-Or, "Moving Shadow Detection by Nonlinear Tone-Mapping," in *Proc. of the 19th International Conference on Systems, Signals and Image Processing (IWSSIP 2012)*, Vienna, 2012. |
| [3] | Y. Moshe and H. Hel-Or, "Video Block Motion Estimation Based on Gray-Code Kernels," *IEEE Trans. on Image Processing,* vol. 18 , no. 10, pp. 2243-2254, 2009. |
| [4] | "Signal and Image Processing Lab.," 2018. [Online]. Available: http://sipl.technion.ac.il/. |

לתעד את הדרכים שניסינו להתמודד עם הבעיה הזו ואתגרים במהלך הדרך

נסינו להשתמש בכל באלגוריתמים הקיימים - מלבד ערכית הפרמטרים שלהם מבחינתנו הם היו כקופסה שחורה ופחות התעסקנו עם המימוש שלהם:

* PyDub VAD: PyDub is a Python library for audio processing, and it provides a simple VAD implementation.
* WebRTC VAD: WebRTC VAD is an open-source VAD library developed by Google. You can use the webrtcvad Python library to integrate it into your projects.
* Spleeter: Spleeter is a pre-trained deep learning model for audio source separation, which can be used for VAD as well.
* Google Cloud Speech-to-Text: Google's Speech-to-Text API can be used for VAD by analyzing the transcriptions it provides. You'll need a Google Cloud account and API key to use this.
* Librosa + Custom Thresholding: You can use the Librosa library for audio analysis and apply custom thresholding for VAD based on the audio's amplitude.
* Kaldi: Kaldi is a popular open-source toolkit for speech recognition, but it also includes VAD functionality. It provides a comprehensive set of tools for speech processing.
* SpeechRecognition:

SpeechRecognition is a widely used Python library that provides a simple interface for working with various speech recognition engines, including Google Web Speech API

* Google Cloud Speech-to-Text:

Google Cloud offers a Speech-to-Text API that allows you to perform speech recognition in the cloud. You can use the google-cloud-speech Python library to interact with this API.

* Naïve power - we tried to separate between the record chunks using measure of the volume highness and to separate each time we measure under certain amount
* **Google VAD -**

Google VAD (Voice Activity Detection) is a Python library that can be used to detect speech in audio signals. It is a two-stage process:

Pre-processing: The audio signal is pre-processed to remove noise and other artifacts. This includes steps such as bandpass filtering, normalization, and framing.

Classification: The pre-processed audio signal is then classified as speech or non-speech using a machine learning model. This model is trained on a dataset of audio signals that have been labeled as speech or non-speech.

The Google VAD library uses the following modules:

* numpy: This module provides a high-performance array data structure and a variety of mathematical functions.
* scipy: This module provides a library of scientific computing tools.
* librosa: This module provides a library for audio analysis.
* tensorflow: This module provides a framework for machine learning.
* Mfcc - MFCC stands for Mel-frequency cepstral coefficients. It is a feature extraction technique that is commonly used in speech recognition and other audio processing applications. MFCCs are derived from the short-term Fourier transform (STFT) of the audio signal. The STFT converts the audio signal into a time-frequency representation. The MFCCs are then calculated from the STFT coefficients.
* Dwt - The Discrete Wavelet Transform (DWT) is a mathematical transform that decomposes a signal into a set of coefficients that represent the signal's frequency and time content. The DWT can be used for a variety of signal processing tasks, such as denoising, compression, and feature extraction.

In Python, the DWT can be implemented using the pywt library. The pywt library provides a number of functions for performing the DWT, as well as for visualizing the results.

* sox - Sox is a command-line tool for manipulating audio files. It can be used to convert between different audio formats, change the speed or pitch of audio, and apply a variety of effects. Sox is a powerful tool that can be used for a variety of audio processing tasks.

PySox is a Python wrapper around Sox. It allows you to use Sox from Python code. PySox provides a number of functions for performing audio processing tasks, such as converting audio formats, changing the speed or pitch of audio, and applying effects.

בסופו של דבר השתמשנו Google VAD:

- עד כמה הטעות משמעותית וכמה נרצה לאפטם את הדאטה

לעדכן את ההקלטות כי אריאל מוסיף עוד פציינטים כל הזמן