

עבוד אותות ושטות למידה לדיבור

תרגיל בית 2 – הפחתת רעשים

בתרגיל זה נבחן שיטות שונות להפחתת רעשים חד ורע ערוצית.

הגדרות

• **סימולציית חדר:** השתמשו בסימולטור `rir_generator` ליצירת חדר בגודל $[4,5,3]$:

<https://github.com/audiolabs/rir-generator>

עם זמני הדהוד: 150ms, 300ms

הגדירו את אורך התגובה לפי: $nsample = T_{60} \cdot f_s$

• **מעריך מיקרופונים:** הגדירו מעריך לינארי אחיד של 5 מיקרופונים לאורך ציר x , עם מרחק של 5 ס"מ בין כל זוג

מיקרופונים סמוכים. מרכז המעריך ממוקם ב- $[2,1,1.7]$.

• **מיקום המקור:** מיקום המקור הוא בזווית של 30 מעלות ובמרחק של 1.5 מטרים ממרכז המעריך.

• **שני סוגי רעשים:**

1. רעש גאוסני לבן מרחבית וספקטראלית.

2. דובר מפריע בזווית של 150 מעלות ובמרחק של 2 מטרים ממרכז המעריך.

$SNR = [0, 10] \text{ dB}$

• **אותות דיבור:** השתמשו באותות מתוך מאגר LibriSpeech בו השתמשתם בתרגיל בית 1.

• **מדדי איכות:** PESQ, ESTOI, SI-SDR – היעזרו בחבילה

[/https://lightning.ai/docs/torchmetrics/stable/gallery/audio](https://lightning.ai/docs/torchmetrics/stable/gallery/audio)

שאלה 1 – יצירת האותות וניתוחם

א. צרו את תגובות החדר המתאימות ושרטטו את הייצוג שלהן במישור הזמן עבור זמני הדהוד שונים

(רק עבור המיקרופון הראשון).

ב. צרו את האותות הנמדדים על ידי מעריך המיקרופונים ע"י פעולת קונבולוציה.

ג. הוסיפו כל אחד מסוגי הרעשים לאות הנמדד (כל פעם רעש אחד).

- ד. שרטטו את האותות הבאים במישור הזמן והתדר (ספקטרום): האות המקורי, האות הנקי הנמדד על ידי המיקרופון הראשון והאות המורעש במיקרופון הראשון. הציגו שרטוטים עבור זמן הדהוד של 300ms ושני סוגי הרעשים בעוצמה של 10dB.
- ה. שמרו כל אחד מהאותות כקובץ wav והאזינו להם.

שאלה 2 – הפחתת רעש רב-ערוצית

- הניחו שמיקרופון הייחוס הוא המיקרופון שממוקם במרכז המערך.
- א. ממשו Delay-&-Sum Beamformer. הניחו שמיקוצי המיקרופונים והדובר ידועים.
- ב. ממשו MVDR Beamformer:
1. שערכו את מטריצת הקווריאנס של הרעש בהתבסס על אותות הרעש.
 2. שערכו את ה-RTF באמצעות שיטת GEVD בהתבסס על אותות הדיבור הרועשים וקווריאנס הרעש המשוערך.
 3. חשבו את משקולות המסנן.
 - ג. הפעילו את המסננים על האותות הרועשים.
 - ו. שרטטו את אותות המוצא במישור הזמן והתדר ושמרו את האותות כקבצי wav. הציגו שרטוטים עבור זמן הדהוד של 300ms ושני סוגי הרעשים בעוצמה של 10dB.
 - ד. חשבו את מדדי האיכות הבאים: PESQ, ESTOI, SI-SDR עם מיצוע על-פני 20 אותות דיבור שונים.
- סכמו בטבלה או בגרף את המדדים עבור שני סוגי הרעשים, עוצמות רעש שונות וזמני הדהוד שונים. שימו לב שאות הייחוס הוא האות הנקי הנקלט במיקרופון הראשון (לא אות המקור).

שאלה 3 – הפחתת שמע באמצעות מודל מבוססי למידה עמוקה חד-ערוצית

- א. הורידו את הקוד של רשת denoiser מבוססת Demucs מתוך הקישור הבא
- <https://github.com/facebookresearch/denoiser>
- בקוד יש מימוש של שתי רשתות (DNS) Deep Noise Suppression וניתן לבחור אחת מהן.
- תוכלו להיעזר בדוגמת ההרצה הבאה <https://colab.research.google.com/drive/1Too3cnMpyKaLQ0vPwDw7jUx0Y3eXm2IA?usp=sharing>
- אין צורך לאמן את הרשת. כחלק מהדוגמה המצורפת ב-colab, עבור מודל מאומן, נעשה ניסיון

להוריד ולטעון את סט המשקולות. במידה ופעולת ההורדה נכשלת, ניתן להוריד אותן ידנית באופן הבא -

עבור רשת dns48

<https://dl.fbaipublicfiles.com/adiyoss/denoiser/dns48-11decc9d8e3f0998.th>

עבור רשת dns64

<https://dl.fbaipublicfiles.com/adiyoss/denoiser/dns64-a7761ff99a7d5bb6.th>

בנוסף, בדוגמת ה-colab המצורפת משתמשים בפקודת `torchaudio.load` על מנת לטעון את קבצי השמע. במידה והפעולה נכשלת, ניתן להשתמש `soundfile.read` (או כל ספרייה אחרת לטעמכם).

ב. הפעילו את המודל על האות המורעש הנקלט במיקרופון הראשון.

ג. חשבו את מדדי האיכות הבאים: PESQ, ESTOI, SI-SDR על אותו סט של 20 אותות דיבור שונים, והשוו לביצועים שהתקבלו בשאלה 2.