

עובד אוטות ויטוות למידה לדיבור

תרגיל בית 2 – הפחחת רעשים

בתרגיל זה נבחן שיטות שונות להפחחת רעשים חד ורב ערכיות.

הגדרות

- **סימולציה חדר:** השתמשו בסימולטור rir_generator [4,5,3] :

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

עם זמני הדחוד: 150ms, 300ms

הגדירו את אורך התגובה לפ' $f_s \cdot nsample = T_{60}$

- **מערך מיקרופונים:** הגדרו מערך לינארי אחד של 5 מיקרופונים לאורך ציר x, עם מרחק של 5 ס"מ בין כל זוג מיקרופונים סמוכים. מרכזו המערך ממוקם ב- [2,1,1.7].

- **מקום המקור:** מקום המקור הוא בזווית של 30 מעלות ובמרחק של 1.5 מטרים ממרכז המערך.

שני סוגי רעשים:

1. רעש גאוסי לבן מרחבית וספקטראלית.

2. דובר מפרייז בזווית של 150 מעלות ובמרחק של 2 מטרים ממרכז המערך.

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

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

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

<https://lightning.ai/docs/torchmetrics/stable/gallery/audio>

שאלה 1 – ייצירת האותות וניתוחם

א. צרו את תגובות החדר המתאימות וشرطו את הייצוג שלهن במישור הזמן עברו זמני הדחוד שונים

(רק עברו המיקרופון הראשון).

ב. צרו את האותות הנמדדים על ידי מערך המיקרופונים ע"י פועלות 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 עברו זמן הדוחה של 300ms ושני סוגי הרעשים בעוצמה של 10dB.
- ה. שמרו כל אחד מהאותות כקובץ wav והאזינו להם.

שאלה 2 – הפחיתת רעש רב-ערוצית

- הניבו שמיקרופון היחסן הוא המיקרופון שממוקם במרכזו המערך.
- משו Delay-&-Sum Beamformer. הניבו שמיוקצי המיקרופונים והדובר ידועים.
 - משו MVDR Beamformer:
 - שרכו את מטריצת הקווריאנס של הרעש בתביסס על אותן הרעש.
 - שרכו את RTF באמצעות שיטת GEVD בהתבסס על אותן הדיבור הרועשים וקווריאנס המושער.
 - חשבו את משקלות המسنן.
 - הפעילו את המسنנים על אותן הרועשים.
 - שרטו את המוצא במישור הזמן והתדר ושמרו את אותן כקובץ wav. הציגו שרטוטים עברו זמן הדוחה של 300ms ושני סוגי הרעשים בעוצמה של 10dB.
 - חשבו את מדדי האיכות הבאים: PESQ, ESTOI, SI-SDR עם מיצוע על-פני 20 אותן דיבור שונים.

כמו בבלה או בגרף את המדים עברו שני סוגי הרועשים, עצומות רעש שונות וזמן הדוחה שונים. שימו לב שוות היחסן הוא האות המקורי הנקלט במיקרופון הראשון (לא אות המקורי).

שאלה 3 – הפחיתת שמע באמצעות מודל מבוסטי למידה عمוקה חד-ערוצי

- הורידו את הקוד של רשות denoiser מבוססת מתוך הקישור הבא
<https://github.com/facebookresearch/denoiser>
 בקוד יש מימוש של שתי רשתות (DNS) וניתן לבחור אחת מהן.
<https://colab.research.google.com/drive/> [1Too3cnMpyKaLQ0vPwDw7jUx0Y3eXm2IA?usp=sharing](https://colab.research.google.com/drive/1Too3cnMpyKaLQ0vPwDw7jUx0Y3eXm2IA?usp=sharing)

אין צורך לאמן את הרשות. כחלק מהדוגמא המצורפת ב-colab, עברו מודל מאומן, נעשה ניסיון

להוריד ולטען את סט המשקولات. במידה ופעולת ההורדה נכשلت, ניתן להוריד אותו 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, בדוגמה ה- colab המצורפת משתמשים בפקודת `torchaudio.load` על מנת לטעון את קבצי השמע. במידה והפעולה נכשلت, ניתן להשתמש `soundfile.read` (או כל ספרייה אחרת לטעמכם).

- ב. הפעילו את המודל על האות המורעש הנקלט בマイיקרופון הראשון.
- ג. חשבו את מדדי האיכות הבאים: PESQ, ESTOI, SI-SDR על אותו סט של 20 אותן דיבורים שונים, והשו **לbijouteries שהתקבלו בשאלת 2.**