**EIE558 Speech Processing and Recognition**

**Lab1: DNN-Based Speech Enhancement**

**A. Objectives and Outcomes**

After finishing this lab, you should be able to perform the following

* Use Google Colab to perform speech enhancement
* Use PyTorch to implement a U-Net for speech enhancement
* Understand the principle of DNN-based speech enhancement

**B. Google Colab**

Google Colab is a free cloud service with GPU support. You may use Colab to develop deep learning applications based on popular libraries such as Keras, TensorFlow, PyTorch, and OpenCV.

**C. Submission**

Write a report, convert it to PDF, and submit it to Blackboard before the deadline. Your report may contain the following:

1. Discussions on your observations, e.g., the correlation between PESQ scores and perceptual quality of listening tests.
2. Waveforms and spectrograms of clean, noisy, and denoised speech
3. How the U-net performs speech enhancement
4. A plot showing the loss against the number of processed mini-batches of the first epoch (this is optional as it takes several hours to train a network on Colab)
5. The effect of removing the skipped connections in a U-Net
6. The effect of reducing the number of frequency bins in FFT (N\_FFT)
7. The effect of reducing the number of feature maps in each convolutional layer
8. The effect of some other modifications that you can think of

**D. Procedures**

***D.1 Prepare Colab Environment***

1. Colab runs on browsers. You need a Google account to use Colab. If you do not have one, visit https://support.google.com/mail/answer/56256?hl=en.
2. Display the Google Drive page (https://drive.google.com/drive/my-drive) in your browser. Use the “+ New” button on the left panel to create the following directory structure in your Google Drive:

My Drive/Learning/EIE558/

After creating the folders, you should see something like this:

Graphical user interface, text, application, chat or text message

Description automatically generated

Directory structure in Google Drive

Later, you may mount your folders in Google Drive to the working environment of Colab.

1. Go to [http://bioinfo.eie.polyu.edu.hk/download/EIE558/DCUnet](http://bioinfo.eie.polyu.edu.hk/download/EIE558/DCUnet%20) and download the file “DCUnet16.ipynb”.[[1]](#footnote-1) Save the file as “DCUnet16.ipynb” on your Desktop. Upload the file to your Google Drive under “My Drive/Learning/EIE558” and remove the added extension “.txt” if your browser adds one for you.
2. Right click the file “DCUnet16.ipynb” on your Google Drive folder and then select “Open with”. If you see “Google Colaboratory”, select it. If your Google account has not been connected with Colaboratory yet, select “Connect more apps” and then select Colaboratory. Then, click “Install”. Then, right click the file again and select “Google Colaboratory”.

Graphical user interface, text, application, chat or text message

Description automatically generated

1. Configure the Colab to use GPU by clicking Edit 🡪 Notebook settings. Select “GPU” in the pop-up window:

Graphical user interface, text, application, email

Description automatically generated

1. Follow the procedure in this Colab file. Note that you may use the pre-trained models in the model/ folder to avoid the lengthen training process.

***D.2 Test on More Noisy Wave Files***

1. Download extra noisy wave files with different signal-to-noise ratio (SNR) from:

<https://polyuit-my.sharepoint.com/:f:/g/personal/enmwmak_polyu_edu_hk/Elw-kGZVexlGhhMO9nyhymQB9qFm55i4Pz3Uf6o683hRmg?e=mktgRI>

1. Create a new directory named “my\_test\_wavfiles” on your Google drive. Then, upload the noised wave files to this directory. For example:

Graphical user interface, application, table

Description automatically generated

Graphical user interface, application

Description automatically generated with medium confidence

1. Run “dc-unet\_test.py” by specifying the path to the noised wave file and the path to the resulting denoised wave file. For example:

!python3 dc-unet\_test.py --noised\_wav data/my\_test\_wavfiles/Gussian/15dB/sp30.wav -–denoised\_wav audio/Gussian\_15dB\_sp30\_denoised.wav

Try different noisy wave files to see the performance of DCUnet and DCnet.

**References:**

1. <https://medium.com/deep-learning-turkey/google-colab-free-gpu-tutorial-e113627b9f5d>
2. H.S. Choi et al., “The algorithm and code are based on the paper "Phase-Aware Speech Enhancement with Deep Complex U-Net" (2019)

-- END --

1. Alternatively, you may download the GitHub repository <https://github.com/enmwmak/Speech-Enhancement> [↑](#footnote-ref-1)