Data Description for 2019 FEMH Voice Data Challenge

1. **Dataset Summary**

The dataset consists of training data and testing data. Details about the dataset for training data are as follows:

TABLE 1. Demographics of the 200 Functional dysphonia and Pathological Voice Samples.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Number | | Mean Age (y) | | Age range (y) | | Standard Deviation | |
|  | M | F | M | F | M | F | M | F |
| FD | 14 | 36 | 37.1 | 35.7 | 20-60 | 3-70 | 13.2 | 14.9 |
| Pathological | 96 | 54 | 55.4 | 55.7 | 27-84 | 21-83 | 13.0 | 12.5 |

*Abbreviations*: M, male; F, female; FD, Functional dysphonia.

TABLE 2. Disease Categories of the 150 Pathological Voice Samples.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Neoplasm | Phonotrauma | Vocal Palsy | Functional Dysphonia |
| M | 45 | 11 | 40 | 14 |
| F | 5 | 39 | 10 | 36 |

*Abbreviations*: M, male; F, female.

The released dataset contains voice samples and medical record obtained from a voice clinic in a tertiary teaching hospital (Far Eastern Memorial Hospital, FEMH, Taipei, Taiwan). The training dataset includes 50 voice samples of functional dysphonia (defined as active symptoms without structural vocal fold disease) and 150 samples of common voice disorders, including phonotrauma (i.e. vocal nodules, polyps, and cysts), laryngeal neoplasm and unilateral vocal paralysis. Medical record include age, gender, jobs, habits and symptoms, when the voice is the worst, how did it happen, whether experienced previous surgery or not, gastroesophageal reflux, voice questionnaire, etc.. Each person has 33 dimensions of medical records and produces a matrix of size 1\*33.

There are 200 testing samples without any labels for performance evaluation. The completion teams are allowed to use a public voice disorder database from the Massachusetts Eye and Ear Infirmary (MEEI) [1] to enhance model and increase samples while taking the risk of possible impact of database mismatch.

All FEMH voice samples are 3-s sustained vowel sound /a:/, which were recorded at a comfortable level of loudness, with a microphone-to-mouth distance of approximately 15*−*20 cm, using a high-quality microphone (Model: SM58, SHURE) with a digital amplifier (Model: X2u, SHURE). The sampling rate was 44100 Hz with a 16-bit resolution, and data were saved in an uncompressed .wav format.

1. **Description of the excel file**

The detail of the .xlsx file is in Table 3 and 4：

TABLE 3. The content of each column.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Folder path |  | | | | | |
| File name | Diagnosis | The patient answered 33 questions | | | | |
| 001.wav  002.wav  …… | 0 : Functional dysphonia  1 : Neoplasm  2 : Phonotrauma\*  3 : Vocal palsy | Gender | Age | …….. | VHI-10 | RSI |
| 1 : Male  2 : Female | The age of patients | …….. |  |  |

\*: include vocal nodules, vocal polyps, and vocal cysts.

TABLE 4. The 33 items of demographics and symptomatic features.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Items | Respond | Feature representation | FD  (n = 50) | Neoplasm  (n = 50) | Phonotrauma  (n = 50) | Vocal Palsy  (n = 50) |
| Sex | male / female | 1/2 | 14/36 | 45/5 | 11/39 | 40/10 |
| Age | The age of patients |  | 36.96±13.5 | 62.8±12.39 | 41±12.78 | 59.1±13.78 |
| Duration (Y) | Sick period(year) |  | 2.58±3.37 | 0.82±1.38 | 0.52±1.25 | 0.22±0.67 |
| Duration (M) | Sick period(month) |  | 1.36±2.27 | 2.1±2.41 | 2.46±2.09 | 3.04±2.22 |
| Hoarseness(Husky voice) | no / yes | 0/1 | 13/37 | 6/44 | 1/49 | 8/42 |
| Narrow pitch range | no / yes | 0/1 | 18/32 | 31/19 | 20/30 | 39/11 |
| Decreased volume | no / yes | 0/1 | 25/25 | 27/23 | 33/17 | 17/33 |
| Tiredness(Fatigue) | no / yes | 0/1 | 11/39 | 42/8 | 20/30 | 30/20 |
| Dysphonia | no / yes | 0/1 | 35/15 | 33/17 | 26/24 | 33/17 |
| Dryness | no / yes | 0/1 | 13/37 | 30/20 | 18/32 | 35/15 |
| Lumping | no / yes | 0/1 | 22/28 | 37/13 | 36/14 | 43/7 |
| Heartburn | no / yes | 0/1 | 46/4 | 47/3 | 48/2 | 49/1 |
| Night meal | no / yes | 0/1 | 44/6 | 47/3 | 48/2 | 48/2 |
| Choking | no / yes | 0/1 | 43/7 | 45/5 | 45/5 | 25/25 |
| Eyes dryness | no / yes | 0/1 | 31/19 | 42/8 | 36/14 | 46/4 |
| Postnasal drip (PND) | no / yes | 0/1 | 28/22 | 46/4 | 38/12 | 45/5 |
| Diabetes | no / yes | 0/1 | 48/2 | 45/5 | 42/8 | 46/4 |
| Hypertension | no / yes | 0/1 | 48/2 | 34/16 | 47/3 | 43/7 |
| Coronary artery disease (CAD) | no / yes | 0/1 | 50/0 | 48/2 | 50/0 | 46/4 |
| Head and neck cancer | no / yes | 0/1 | 50/0 | 48/2 | 50/0 | 44/6 |
| Head injury | no / yes | 0/1 | 50/0 | 50/0 | 50/0 | 48/2 |
| Cerebral vascular accident (CVA) | no / yes | 0/1 | 50/0 | 50/0 | 50/0 | 49/1 |
| Smoking | never/past/active | 0/1/2 | 45/3/2 | 13/12/25 | 44/1/5 | 29/15/6 |
| PPD(pack of cigarette per day) |  | 0/0.5 to1/1.5 to5 | 48/2/0 | 25/18/7 | 45/5/0 | 45/4/1 |
| Alcohol Drinking | never/past/active | 0/1/2 | 31/0/19 | 31/1/18 | 34/0/16 | 40/1/9 |
| Drinking frequency | not/occasionally/weekly/daily | 0/1/2/3 | 31/17/1/1 | 30/16/3/1 | 34/15/0/1 | 42/6/0/2 |
| Noise at work | not/a little/noisy | 1/2/3 | 30/17/3 | 31/12/7 | 16/18/16 | 38/10/2 |
| VAS | worst to best | 0 to 10 | 3.861.50 | 2.61.79 | 2.681.43 | 2.241.57 |
| Voice handicap index – 10 | Sum of 10-item voice handicap index | 0 to 40 | 21.726.77 | 24.446.77 | 24.6.54 | 30.667.33 |
| Reflux symptom index | Sum of 9-item reflux symptom index | 0 to 45 | 14.647.12 | 10.926.71 | 13.126.76 | 16.549.59 |
| Onset of dysphonia | sudden  gradually  on and off  since childhood  other | 1  2  3  4  5 | 7  22  5  8  8 | 15  24  8  0  3 | 25  15  10  0  0 | 25  14  0  0  11 |
| Diurnal patterns | worse in the morning  worse in the afternoon  similar all day  fluctuating | 1  2  3  4 | 11  7  12  20 | 5  11  22  12 | 13  18  7  12 | 2  9  30  9 |
| Occupational Vocal demand | always  frequent  occasional  minimal | 1  2  3  4 | 24  15  10  1 | 7  14  11  18 | 27  15  5  3 | 12  16  10  12 |

*Abbreviations*: FD, Functional dysphonia.