Brief Instruction on Analyzing the BabyEar4k Dataset

Liu-Jie Ren, [renliujie@fudan.edu.cn](mailto:renliujie@fudan.edu.cn)

Fudan University

This is a brief instruction on analyzing the BabyEar4k dataset, which contains 3,852 ear images from newborns, all with a diagnostic result of auricular deformities. The healthy data of the babies and their mothers are also provided.

The dataset analysis mainly includes the following two aspects: (1) calculating the average ears, (2) correlation analysis of auricular deformities.

# Software Environment Requirement

The following analysis was conducted in Windows 10 64 bit, Matlab R2020a. However, the code should also work on other platforms (Windows 11, Linux) and other Matlab versions (R2020a and later versions).

# Preparing

The dataset is included in the ‘./BabyEar4k/’ directory.

* ‘./BabyEar4k/images’ is a folder containing all ear images.
* ‘./BabyEar4k/diagnosis\_result.csv’ is the diagnosis result file.
* ‘./BabyEar4k/health\_data.csv’ contains the health information.
* ‘./BabyEar4k/bounding.csv’ contains the bounding annotations of the ears

Please put all the code files (\*.m) should be in the same directory with ‘./BabyEar4k’.

# Code files

Three code files were programmed, they are

* *BabyEar4k.m*: this is a Matlab class for coping with the BabyEar4k dataset, the class contains the methods for data reading and analyzing.
* *averageEars.m*: this is a Matlab code that generate the average ears.
* *correlationLeftRight.m*: this is a Matlab code that analyzing the correlation between left and right ear deformities.
* *correlationHealth.m*: this is a Matlab code that analyzing the correlations between the health data and ear deformities.
* *chi2test.m*: this is a Matlab function for chi-square test.

# Usage

* to calculating the average ears, the users can simply run the code *averageEars.m*
* to analyze the correlations between the left and right ears, just run *correlationLeftRight.m*
* to analyze the correlations between the ear deformities and health data, run *correlationHealth.m*
* Note: feel free to revise all the codes for your own purposes.