Ministère de l'Enseignement Supérieur
et de la Recherche Scientifique

Université de Carthage

Ecole Polytechnique de Tunisie

وزارة التعليم العالي و البحث العلمي

جامعة قرطاج

المدرسة التونسية للتقنيات

Delivered Report

Multibranch Deep Neural Networks for the Detection of Vocal Cord Pathologies on Videos

13/07/2020 - 21/08/2020

Host Laboratory:

COSIM
Communications, Signal and Image

Elaborated by:

Eya Ghamgui
Third Year Engineering Student

Supervisors:

Amel Ben Azza
Professor at Carthage University & Researcher at COSIM lab/ Sup'Com

Dhouha Attia
Associate professor at Carthage University & Researcher at COSIM lab/ Sup'Com

Academic Year
2020/2021

Contents

Image Modifications

1.1 Edge Detection

The code **Edges** is a matlab code. It takes as input binary images where the ground truth is found have the value 0 and the pixels of the glottal area have the value 255. This code uses morphological, seedpoint, and segCroissRegion functions developed by Abdelahmid. Finally, it returns images of the corresponding edges of the glottal area.

1.2 Data Division

Each notebook **(i.e. TransferImg_train, TransferImg_val, and TransferImg_test)** is used to divide the initial data into folders grouping the images per patient.

1.3 Data Re_indexing

The **photo_index** notebook is used to index the initial data in an increasing order per cycle.

1.4 Create Complete Cycles

In the notebook **Decompose Data**, we developed a code that creates a new dataset of images for all the complete cycles. It takes as input the initial images and the INFO_subjects _out excel file. After that, the code creates new folders, 2 folders for each cycle. The first one contains the sequence of images of the corresponding cycle and the second one contains its corresponding ground truth.

1.5 Counters

The **Contour** notebook draws the counters on the initial images. It takes as input the images of the binary counter and the initial images and it returns ones with the drawn green counters.

Histograms

The notebook **Histograms_27_07_2020** is used to draw the histograms of the duration of cycles and the plots of the mean of the GAW. It takes as input the INFO_subjects excel file and the gaw csv file. The first file contains all the information about the subjects and their corresponding cycles, with their duration, state and class. After making some changes on this dataset, we added a code to save the new one in a new excel file intitled INFO_subjects_out. The second file contains the GAW vectors for each complete cycle. In this notebook, We used different python libraries such as matplotlib and seaborn to have better plot and the numpy and pandas to deal with datasets.

GAW Features

The code **GAW_Features** is a matlab code. It calculates all the GAW features using the ratios function. The Ratio_matrix determines the features calculated by Abdelahmid. The Ratio_matrix1 contains the values of glottal gap index of the Left and Right Vocal Folds (GGI), amplitude to length ratio of the left and right vocal folds (ALR), closing quotient of the left and the right vocal folds (CQ), and the stifness (ST). Finally, this code saves all this values in a csv file named ratios.

Segmentation

4.1 Threshold Method

The code **Threshold_Method** is a matlab code. It uses the same functions defined by Abdelahmid in his internship in order to perform the threshold segmentation method but applied our dataset. In other words, we applied the main function on each complete cycle. In addition, we considered a complete cycle as a video sequence.

4.2 Adjusted Mean Shift

The **Adjusted Mean Shift** notebook presents the mean shift segmentation technique using the ROI detection method. It takes as input the images of all the complete cycles and returns their segmentation results.

4.3 DL Segmentation

The notebook **Compute_performance_on_test_set** takes as input all the complete cycles images and their binary masks, the Utils folder, Unet8 pre-trained model and the INFO_subjects_out excel file. It returns the deep learning segmentation results of the UNet model developed by Ronneberger et al. and pre-traned on the BAGLS dataset.

4.4 Performance Metric

In the notebook **Performance_Metrics**, we calculated the segmentation performance metrics for the three segmentation techniques: the threshold method, the mean shift method and the deep learning method. After that, we saved all the results obtained in excel files. In addition, we plotted all the performance metrics for all cycles. Finally, we drew the bar plot of the mean of all cycles of this performance metrics. This notebook takes as input the segmented images for each segmentation technique.

Classification

5.1 SVM

In this part, we developed three notebooks for each method: SVM applied on the GAW features, SVM using the PCA method and SVM using the ReliefF method. We created the same notebook for the feature calculated on the ground truth images and the deep learning segmentation results.
Each notebook takes as input INFO_subjects_out excel file and ratios csv file. Using these information we developed binary and multiclass SVM. We studied the balanced binary SVM for the healthy and the unhealthy classes, imbalanced binary SVM for the healthy and the unhealthy classes and imbalanced binary SVM for the healthy and diseases classes. For the multiclass SVM, we studied the oneVsone and oneVsrest methods and we added the calibration technique for the oneVsrest method. In each part, we drew the decomposition of the datasets as a pie chart. Also, we calculated classification metrics (f_1 score, auc, accuracy, precision and recall) and we drew the corresponding confusion matrices. In addition, we performed the cross validation technique. In this step, we created the folds manually by dividing our datasets into 5 equaly folds and we repeated the same previous work.

5.2 The ResNet

This notebook (**ResNet**) develops an end-to-end deep learning model that classify vocal folds diseases. This model uses the transfer learning technique and the ResNet50 as a pre-trained model. The notebook takes as input resnet50_weights_tf_dim_ordering_tf_kernels.h5

file which is the pre-trained model and INFO_subjects_out excel file which contains all the information about the complete cycles and also all the complete cycles images. After restoring the pre-trained, we added dense layers along with activation and batch normalization. Finally, we drew the corresponding confusion matrices.