Tutorial 11: EAT – The ICMI 2018 Eating Analysis and Tracking Challenge (40P + XP)

Please submit your predictions, code and presentation by 12 July 14:15 to manuel.milling@informatik.uni-augsburg.de AND maurice.gerczuk@informatik.uni-augsburg.de.

The EAT Challenge

This year's deep learning tutorial challenge will be based on the sub-task **Food Type** of the ICMI 2018 Eating Analysis and Tracking Challenge. It is an audio-based task, in which speech recordings have to be classified based on the type of food the speakers are eating while the recordings are being made. For this challenge we will only use the audio data. For further information on the challenge as well as some baseline approaches please refer to the *paper on digicampus*.

Get Access to Data

Follow the steps below.

- Download and sign (digital signature is okay) the End User Licence Agreement. Send the signed document to manuel.milling@informatik.uni-augsburg.de in order to receive the password for the data.
- Download the `EAT_data.zip` from `https://megastore.uni-augsburg.de/get/rXvJzBqyAJ/` and unzip it using the password from the previous step.

Data Structure

The data contains the following files and directories:

- `README_Test`: Read this file carefully, especially 'section 1): Submission format'. Your submission needs to be exactly of the form specified in **Option 1**. *Note: Be reminded that we only consider the food type sub-task like in the given example (classes: Crisp, Nectarine, Haribo, etc.)*
- `README`: This file specifies how to use the code of the provided baselines and might therefore be helpful but not necessary for your solution.
- `audio/`: This directory contains the raw audio files for the training and test partition of the form `train_XXX.wav`. Samples are usually around 5-15 s long with a sampling rate of 44,1 kHz.

- `audio_features/`. In this directory you can find a file for each the training and test directory containing audio COMPARE features for each file and 10 ms. *Note: This directory is useful if you plan to use said COMPARE features in your approach.*
- `audio_features_instance/`: This directory contains COMPARE features as the one mentioned above but only one per audio sample.
- `labels/`: The labels for the training as well as the metadata for the test data. *Note: Be reminded that we only consider the food type sub-task like, i.e., class labels like Crisp, Nectarine, Haribo, etc.).*
- `baseline_xbow/`, `baseline_end2you/`: Baseline scripts.

1 Challenge Participation (30P)

Train your models based on the training data and submit up to three predictions for the test data as specified in the `README_Test`, as well as your code. Different predictions can be based on different approaches or for instance the same approach trained with different random seeds. The predictions have to be exactly of the specified form as we are using automated scripts to evaluate them.
You can use any machine learning approach, as long as it has at least one deep learning component and must not be identical to the baselines provided or published approaches on the dataset (though inspiration from literature or the baselines is highly welcome). Examples are the Spectrogram-CNN approach used for exercise sheet 07, the end2you baseline of the challenge or RNN-based approaches utilising extracted audio features. Some further information and tools will be presented in the tutorial.
The test labels may only be used to create the predictions and of course manual labelling of the test data is not allowed.
Note: Results should be clearly above chance level (i.e., 30% unweighted average recall (UAR) on the test set or above).

2 Bonus: Beat the Baseline(XP)

Beat a baseline for the given sub-task of **49.3 % UAR** on the test data obtained with a spectrogram and deep CNN approach in preparation of the challenge. For each percentage point above 49% you obtain one additional bonus point for the tutorial.

3 Presentation (10P)

Prepare a short presentation (using PowerPoint, LaTeX, or similar) of roughly 3 content slides. The presentation should take about 5 minutes and you should introduce yourself, your approach and your results. You receive your results (UAR of your predictions as well as a confusion matrix) after submitting your predictions and you can submit the presentation after obtaining your results. Present your solution (if possible

as a team) in the final Deep Learning tutorial on 14 July. Please check your camera and microphones on Zoom before the tutorial as the schedule will be quite tight.

Thank you for your hard work and participation this semester :).