ARDUOUS Challenge

Kristina Yordanova Emma L. Tonkin Teodor Stoev
Gregory J. L. Tourte

July 2020

1 ARDUOUS annotation challenge

This challenge focuses on the problem of annotating everyday activities (cooking) captured with camera and sensors. For this purpose we are using the CMU Kitchen Dataset. The challenge has two strands:

- Manual annotation: demonstrate your manual annotation software or toolset using the subset of this dataset described in section 1.1. We will compare the performance via various metrics, including: inter-rater reliability (percentage agreement, Cohen's Kappa, intraclass correlation, Krippendorff's Alpha, Pearson's correlation coefficient, Spearman's Rho), time taken for the annotation process, and learning curve.
- Semi-supervised or automated annotation: build a solution or demonstrate your existing solution for annotating the data from this dataset, using any of the datatypes listed in section 1.1. We will compare the performance with standard machine learning evaluation metrics (accuracy, precision, F1-score). We will also ask you to provide indications regarding the speed and efficiency of the method used, and the resources required to run it.

For both of these tasks we are providing our own data dictionary for you to use, which is more detailed than the existing CMU annotation, and is validated against a series of metrics and against a causal model.

1.1 Dataset description

The CMU Kitchen Dataset is a subset of the CMU Multi-Modal Activity Database (CMU-MMAC), which 'contains multimodal measures of the human activity of subjects performing the tasks involved in cooking and food preparation. The CMU-MMAC database was collected in Carnegie Mellon's Motion Capture Lab. A kitchen was built and to date twenty-five subjects have been recorded cooking five different recipes: brownies, pizza, sandwich, salad, and scrambled eggs.' [1]

The video dataset consists of

[1] http://kitchen.cs.cmu.edu/

1. Three high spatial resolution (1024×768) color video cameras at low temporal resolution (30 Hz).
2. Two low spatial resolution (640×480) color video cameras at high temporal resolution (60 Hz).
3. One wearable high spatial resolution (800×600/1024×768) camera at low temporal resolution (30 Hz).

The audio dataset consists of:

1. Five balanced microphones.

The motion capture dataset consists of:

1. A Vicon motion capture system with 12 infrared MX-40 cameras. Each camera records images at 4 megapixel resolution at 120 Hz.

The internal measurement units (IMUs) dataset consists of:

1. Wired IMUs (3DMGX),
2. Bluetooth IMUs (6DOF).

Finally, the dataset containing data from wearable devices consists of:

1. BodyMedia,
2. eWatch.

For this challenge, we ask you not to use the audio feed or motion capture or the BodyMedia device data. **You can use the video, the eWatch wearable data, the IMUs and RFID data.**

2 Objectives

2.1 Manual annotation approach

For this the dataset is pared-down to a representative subset, since manual annotation of the full set would be highly unreasonable. To download the dataset, run the startup script described in section 2.3 of this file.

The primary goal of the manual annotation challenge is to demonstrate your manual annotation software or toolset. We'd like to see what is innovative about your solution: the process, the methods used, the evaluation approach you use or the interface decisions or approaches that you have taken. We'd also like to know what training you offer to your (human!) annotators.

Demonstrate your manual annotation software or toolset using the runs from the dataset described in table 1. Note, that the required runs are automatically extracted with the help of the script described in section 2.3.

2.1.1 Evaluation metrics for your approach

We will compare the performance via various metrics, including: inter-rater reliability, time taken for the annotation process, and the learning curve evident in the data.

Table 1: Runs to be annotated manually

Recipe	Runs			
Brownies	S54	S32	S31	S09
Sandwich	S16	S25	S34	S15
Eggs	S08	S20	S16	S50

2.1.2 How to submit results

Results should be returned in the format:

```
annotator | datetime | annotation | file annotated | start/end/position in file
```

2.2 Semi-supervised or automated annotation

For this challenge the dataset is separated into *training* and *test* subsets, in an approximate 80–20 split. You can download this dataset using the script described in section 2.3 of this document.

The goal of this challenge is to demonstrate your semi-supervised or automated method for annotating data. We'd like to see your solution and will compare it using the usual machine learning metrics, but also consider the evaluation approach that you propose and the reusability of the solution.

Build a solution or demonstrate your existing solution for annotating the data from the runs listed in table 2 (training data) and table 3 (test data). Note, that the required runs are automatically extracted with the help of the script described in section 2.3.

Table 2: Runs to be used for training

Recipe	Runs			
Brownies	S47	S54	S13	S31
Sandwich	S12	S16	S25	S34
Eggs	S28	S08	S20	S16

Table 3: Runs to be used for testing

Recipe	Runs
Brownies	S09
Sandwich	S15
Eggs	S50

2.2.1 Evaluation metrics for your approach

We ask you to provide evaluation results for your approach. These include precision, recall, accuracy and F1 score. Beyond this we also encourage you to provide further metrics of your choice as various metrics are appropriate depending on your approach.

2.2.2 How to submit results

Results should be returned in the following format:

zip file containing the annotations created on the test set

If relevant for your setup, we would also like you to share the time and resources needed to train the model: please benchmark this with the platform-appropriate equivalent of

```
dir> perf stat training-process.ext
```

2.3 Startup scripts

A startup script is available on Github: https://github.com/etonkin/kitchen_dataset_tools. It downloads the relevant parts of the dataset listed above.

2.4 Data sources

The original source of the datasets used in this challenge are listed in this section:

CMU kitchen dataset:	http://kitchen.cs.cmu.edu/
Rosdock annotations:	http://purl.uni-rostock.de/rosdok/id00002273

2.5 Annotation schema

Details about the annotation provided by the University of Rostock and the underlying semantic models could be found in [1].

The annotation schema used in the provided annotation can be found at https://github.com/etonkin/kitchen_dataset_tools.

The Rostock annotations are based on a very detailed schema, which covers interaction with kitchen storage, appliances and objects used in cooking. The full list of terms used in each recipe represented in the dataset can be found in the Data_Dictionary directory. The annotation follows the schema *action-object/location-object/location*. In case the tool or algorithm you are using is not able to annotate on that granularity level, you are allowed to remove some details from the annotation (e.g. annotating only action classes, or actions and locations, etc.). The level of granularity of your approach will be taken into consideration during the evaluation.

3 Outcomes

A small prize will be offered for the winner in each category. Select participants will be invited to present their results in the next ARDUOUS workshop. The most highly scored overall entry across both challenges will also be granted a fee waiver to present their results in the MDPI journal *Sensors*.

4 Deadlines

The submission deadline is 17th October 2021.

5 Contact

For any questions, contact Kristina Yordanova (kristina.yordanova@uni-rostock.de), Emma Tonkin (e.l.tonkin@bristol.ac.uk), Teodor Stoev (teodor.stoev@uni-rostock.de) or Greg Tourte (g.j.l.tourte@bristol.ac.uk).

6 Acknowledgements

We thank CMU for making this dataset available. We would also like to thank the students who worked on the annotation: Andreas Siebert and Mona Arndt from the University of Rostock. We would also like to thank Frank Krüger together with whom the annotation schema and the process were developed.

References

[1] Yordanova, K. and Krüger, F. 'Creating and Exploring Semantic Annotation for Behaviour Analysis'. In: *Sensors* 18.9 (2018), 2778:1–22. ISSN: 1424-8220. DOI: 10.3390/s18092778.