Using DeepProbLog to perform Complex Event Processing on an Audio Stream

Motivating scenario

- Audio sensors have been deployed in a city.
- We want to detect *complex events*:
  - These are combinations of events distributed in time and space.
  - These can be detected by combinations of sounds defined by an expert.
  - For instance, if there is *shouting*, *glass shattering* and *sirens*, we can detect a *riot*.
- However, we do not have large amounts of training data, as these situations are rare.

Objectives

- Being able to use subsymbolic data as input.
- Retaining flexibility in rule definitions.
- Being able to perform end-to-end training.
- Being robust against noisily labelled data.

Related work and its limitations

There are three main types of approaches to this type of problem:

- Combining pre-trained neural networks with symbolic approaches. [1,2]
  - Requires access to pre-trained neural networks.
- Neural network approaches such as LSTMs [3] or Convolutional 3Ds [4].
  - They require very large amounts of data to train.
  - Does not allow for rule definitions by experts.
- Neuro-symbolic approaches. [5]
  - Currently, they significantly limit rule definitions.

[1] Roldán, J.; Boubeta-Puig, J.; Luis Martínez, J.; and Ortiz, G. 2020. Integrating complex event processing and machine learning: An intelligent architecture for detecting iot security attacks. Expert Systems with Applications 149:113251.

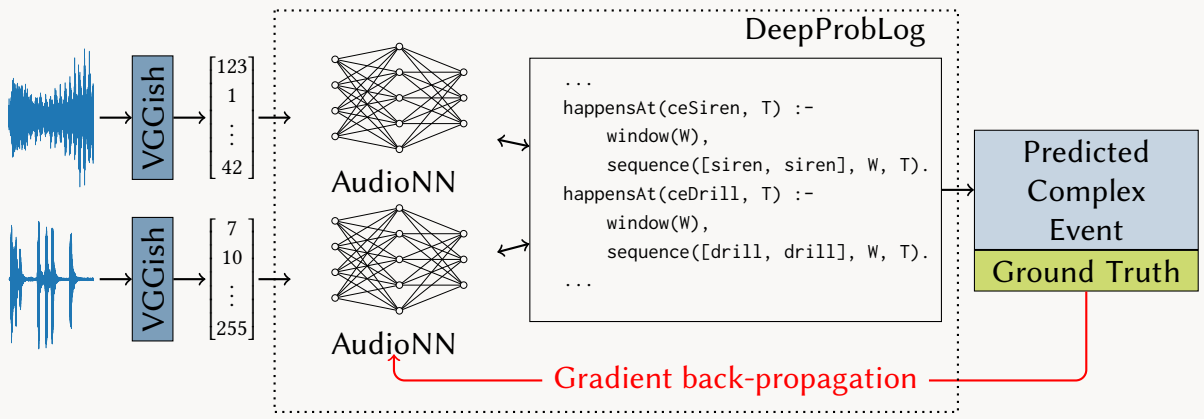
[2] Roig Vilamala, M.; Hiley, L.; Hicks, Y.; Preece, A.; and Cerutti, F. 2019. A pilot study on detecting violence in videos fusing proxy models. In 2019 22th International Conference on Information Fusion (FUSION), 1–8.

[3] Mishra, S.; Jain, M.; Siva Naga Sasank, B.; and Hota, C. 2018. An ingestion based analytics framework for complex event processing engine in internet of things. In Mondal, A.; Gupta, H.; Srivastava, J.; Reddy, P. K.; and Somayajulu, D., eds., Big Data Analytics, 266–281. Cham: Springer International Publishing.

[4] Liu, K.; Liu, W.; Gan, C.; Tan, M.; and Ma, H. 2018. T-c3d: Temporal convolutional 3d network for real-time action recognition.

[5] Xing, T.; Garcia, L.; Vilamala, M. R.; Cerutti, F.; Kaplan, L.; Preece, A.; and Srivastava, M. 2020. Neuroplex: Learning to Detect Complex Events in Sensor Networks through Knowledge Injection. New York, NY, USA: Association for Computing Machinery. 489–502.

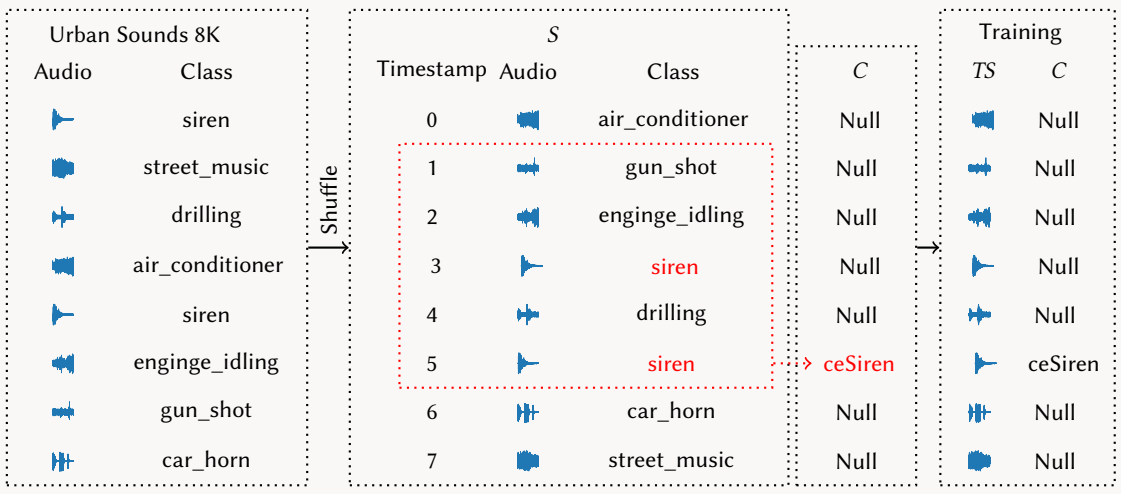
Proposed approach



The ProbLog code in the figure above shows the example rules used in the experiments below. These use the clause *sequence* to define event patterns. Users are able to define more complex rules with the full expressiveness of ProbLog.

Base dataset

We generated a synthetic dataset based on Urban Sounds 8K. The resulting dataset contains 10 types of complex events, which happen when two instances of the same class of sound occur within a window of time. In our experiments, we used window sizes between 2 and 5 seconds.



Random noise datasets

- To evaluate robustness to noisily labelled data, we generated *random noise datasets*.
- These have different percentages of noisy data (0% to 60% of the dataset).
- For this noisy data, the correct label has been replaced by a random one.

Results

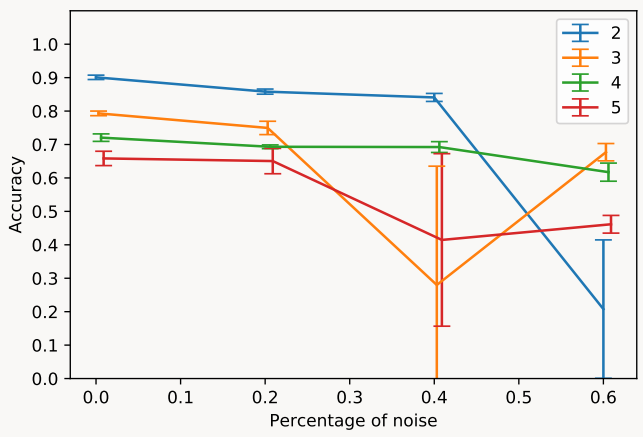


Figure: Evaluation of system's performance with the different percentages of noise from the random noise datasets. Each colour represents a different window size. Error bars show the standard deviation over 3 executions.

- As shown in the figure above, our approach performs fairly well with the base dataset (percentage of noise of 0.0).
- Bigger window sizes perform worse than the smaller ones. This is due to the added complexity of bigger window sizes.
- A percentage of noise of 20% has an almost imperceptible effect on the accuracy of the system.
- Bigger percentages of noise lead to unreliable results, as the system is not able to train correctly in a consistent manner.

Conclusions and future work

- We have designed a system that fulfills all four objectives.
- As future work, we want to consider other types of subsymbolic data.
- Improvements to the time efficiency of our approach would also be a good area of research, as that is a downside of our approach.

Marc Roig Vilamala, Tianwei Xing, Harrison Taylor, Luis Garcia, Mani Srivastava, Lance Kaplan, Alun Preece, Angelika Kimmig, Federico Cerutti

CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD
DAIS ITA