DATA2B

WE MAKE DATA PRODUCTS

Detection of weak signals in video streams with active reinforcement learning

Video digitizes images and audio streams. It is a cost-effective media for remote vision as for sharing captures of physical events occurring around us. For recent uses such as video surveillance, data streams are processed for finding patterns in order to aid the human vision by machine learning. In other domains such as signal processing, tools of observations record phenomena appearing beyond the visual and audio spectrum, producing data that we could compare to video streams. In both cases, some people spend time and life observing data streams for finding special events, being physical risks in sensible places or neutrinos preceding a supernova [3]. As Big Data and uninterrupted data flows easily escape from human attention and understanding, computer vision can massively process these data flows without interruption [4].

For identified events such as fire and flames, computer vision is reliable enough for efficient applications as it can distinguish flames between fire-colored moving objects, thanks to spatial and temporal wavelet transform [5]. Computer vision systems also enable the early detection of identified events and features, for autonomous vehicles and cobots [2]. With unidentified events, the best way to efficiently detect and identify them remains human observation and analysis. As a result, detecting and recognizing unidentified events in data streams requires man-machine collaboration, instead of man-machine competition [6].

Active learning aims at the collaboration of human intelligence and artificial intelligence [7]. With supervised models in machine learning and deep learning, items must be previously labeled by experts and/or people answering to labeling queries – e.g. captcha. Active learning from relative queries, instead of absolute ones, makes possible to train machine-learning algorithms from weak predictors, such as side information or non-experts answers, then to get comparable or better performance than with experts' answers to labeling queries [1]. With active learning for reinforcement learning, items recognition relies on human information processing-based inputs, more than on human knowledge-based inputs [8].

The IoT and smart cities are going to produce huge and heterogeneous datasets, to be compared with astronomic observations and collected in data lakes. Therefore, the detection of unexpected and/or unidentified events in data streams should become a significant task, depending on active learning for the future of society, science, and economy [9-12]. Finally, detecting weak signals in videos streams thanks to active learning or other approaches could represent an important stake, while there currently appears no or few works published on this topic, to our knowledge. It might allow catching unusual phenomenon in video surveillance before specific events occur, so as to prevent risks. It might allow civil security, medical practitioners, seismologists, meteorologists and other professions processing signals to be noticed of unobserved phenomena that could require attention, in order to identify and classify them for an active reinforcement learning process.

The proposed thesis concerns the research and experimentation of an online model of active learning for the detection of weak signals in data streams. It is focused on application in video surveillance and other data streams online applications. It will be experimented thanks to Big Data systems and Lambda architecture hosted and supported within the company DATA2B, allowing to run offline and online analysis tasks in parallel processes. The PhD student will work with private and academic research teams. He/She will be employed by the company (CDD CIFRE). Candidates wishing a permanent position as an associated R&D Director for online analysis after the thesis are welcome.

References

[1] Qian B., Wang X., Wang F., Li H., Ye J., and Davidson I. Active learning from relative queries. In Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence, IJCAI '13, pages 1614–1620. AAAI Press, 2013.

[2] Donadio F., Frejaville J., Larnier S., and Vetault S. Human-robot collaboration to perform aircraft inspection in working environment. In Proceedings of 5th International conference on Machine Control and Guidance,, 2016.

[3] Antonioli P., Tresch Fienberg R., Fleurot F., Fukuda Y., Fulgione W., Habig A., Heise J., McDonald A.B., Mills C., Namba T., Robinson L.J., Scholberg K., Schwendener M., Sinnott R.W., Stacey B., Suzuki Y., Tafirout B., Vigorito C., Viren B., Virtue C., and Zichichi A. Snews: the supernova early warning system. New Journal of Physics, 6(1):114, 2004.

[4] Najafabadi, M. M., Villanustre, F., Khoshgoftaar, T. M., Seliya, N., Wald, R., and Muharemagic, E. Deep learning applications and challenges in big data analytics. Journal of Big Data, 2(1), 1, 2015.

[5] Dedeoglu Y. Güdükbay U. and Cetin A. E Töreyin, B. U. Computer vision based method for real-time fire and flame detection. 1(27):49–58, 2006.

[6] Zheng, N. N., Liu, Z. Y., Ren, P. J., Ma, Y. Q., Chen, S. T., Yu, S. Y., and Wang, F. Y. Hybrid-augmented intelligence: collaboration and cognition. Frontiers of Information Technology & Electronic Engineering, 18(2): 153-179, 2017.

[7] Jordan, M. I., and Mitchell, T. M. Machine learning: Trends, perspectives, and prospects. Science, 349(6245): 255-260. 2015.

[8] Vijayanarasimhan, S., and Grauman, K. Large-scale live active learning: Training object detectors with crawled data and crowds. International Journal of Computer Vision, 108(1-2): 97-114. 2014.

[9] Sharma, M., Das, K., Bilgic, M., Matthews, B., Nielsen, D., and Oza, N. Active learning with rationales for identifying operationally significant anomalies in aviation. In Joint European Conference on Machine Learning and Knowledge Discovery in Databases, 209-225. 2016.

[10] Persello, C., and Bruzzone, L. Active and semisupervised learning for the classification of remote sensing images. IEEE Transactions on Geoscience and Remote Sensing, 52(11): 6937-6956. 2014.

[11] Nguyen, D. H., and Patrick, J. D. Supervised machine learning and active learning in classification of radiology reports. Journal of the American Medical Informatics Association, 21(5): 893-901. 2014.

[12] Rubens, N., Elahi, M., Sugiyama, M., and Kaplan, D. Active learning in recommender systems. In Recommender systems handbook, 809-846. 2015.

Requirements

The candidate must have an MSc in engineering, artificial intelligence, computer science, or a closely related field, with verifiable experience in signal processing and computer science. Intermediate or advanced programming skills are necessary, as well as verifiable speaking and writing skills in English (C1 or C2 level).

Application

Candidates must send all the following documents:

- Detailed Curriculum Vitae
- A motivation letter explaining the interest in the PhD subject
- A summary of the MSc internship
- Copies of engineering and MSc degrees
- Transcripts of the engineering and MSc degrees grades (for foreign students).

to this address:
doctorat-cifre-data2b-imt@imt-atlantique.fr