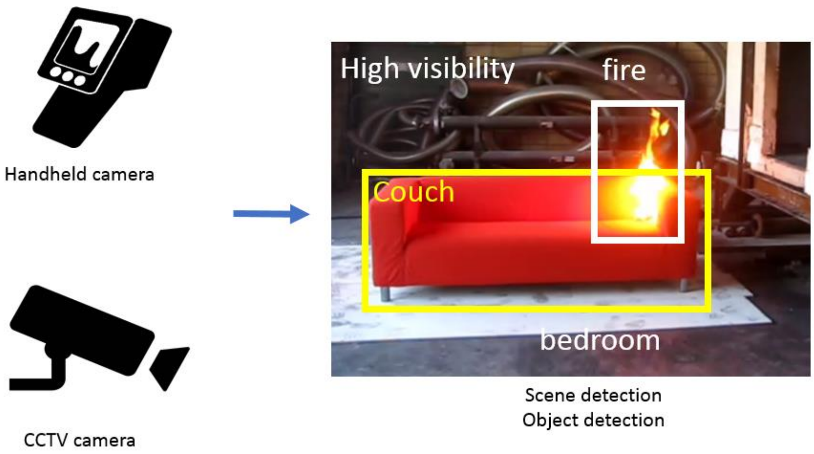

Figure 2. Automated scene content understanding from video footages (e.g., handheld and CCTV cameras).

4.3 Semantic based querying

The web-based tool that could be used to explore the enriched (i.e., the annotated) keyframes is a metadata filtering and clustering service. The querying mechanism can be activated in three different ways:

- A textual input where the user can type the tag that he wants. In case the input is not in the list of predefined tags, the tag with the closest path distance (i.e., the highest similarity) according to Wordnet [28] is selected. However, other semantic distance metrics can easily be integrated.
- A drop-down list with predefined tags (e.g., places, fire-related topics, evacuation keywords), where the user can select the necessary tags (e.g., show all the keyframes annotated with kitchen and fire).
- An interactive, hierarchical ontology visualization where the user can select and easily search for the best matching tags. Furthermore, if a tag on a higher semantic level in the hierarchy is selected, all the underlaying and related tags are selected (e.g., if the tag opening is selected, the tags door and window are likewise selected).

An ontology is a representation, formal naming, and definition of the categories, properties, and the relations between the concepts, data, and entities over one or more domains.

The semantic similarity is currently calculated based on the ontology of Zhang et al. [29], but new ontologies can easily be integrated. Zhang et al. proposed a street scene ontology for qualitative understanding of outdoor scenes. This is valuable for large multi-disciplinary fire incidents as it contains building elements (e.g., floor, window, wall, column), but also construction, land and terrain elements. Some alternative ontologies that can be used are, for example, the work of Kadar et al. [30] who created a database of 100 scene categories (e.g., classroom, bathroom, bedroom, alley) derived from human vision. This ontology could be valuable in indoor firefighting scenarios. Jaoa et al. [31] created an ontology on how scene situations progress in time. The FIRE ontology was created in order to represent the set of concepts about the fire occurring in natural vegetation, its characteristics, causes and effects. Similar concepts and effects are found in indoor fire situations. Poveda et al. [32] created an ontology for designing and validating emergency plans and the sensor, users and furniture connections are highly valuable for our framework.

5 Results and discussion

Murugan et al. [33] proposed various methods for video summarization for surveillance applications still latter discussion was given on the exportability of the video footages. Within this Section a subjective evaluation is performed on our proposed mechanisms. Video footages from multidisciplinary incidents are taken as input for our evaluation. Figure 3 gives a schematic overview. First the video footages are selected (currently, this is a manual task, but the integration of online IP-cameras should be easily possible). Secondly, the keyframe extraction, similarity removal and no-reference analysis is used to select the most representative keyframes. Thirdly, the semantic tag understanding process is elaborated to automatically generate tags for each frame. Finally, the exploration and selection tools are used to get a fast overview of the current state of the incident and the actions.

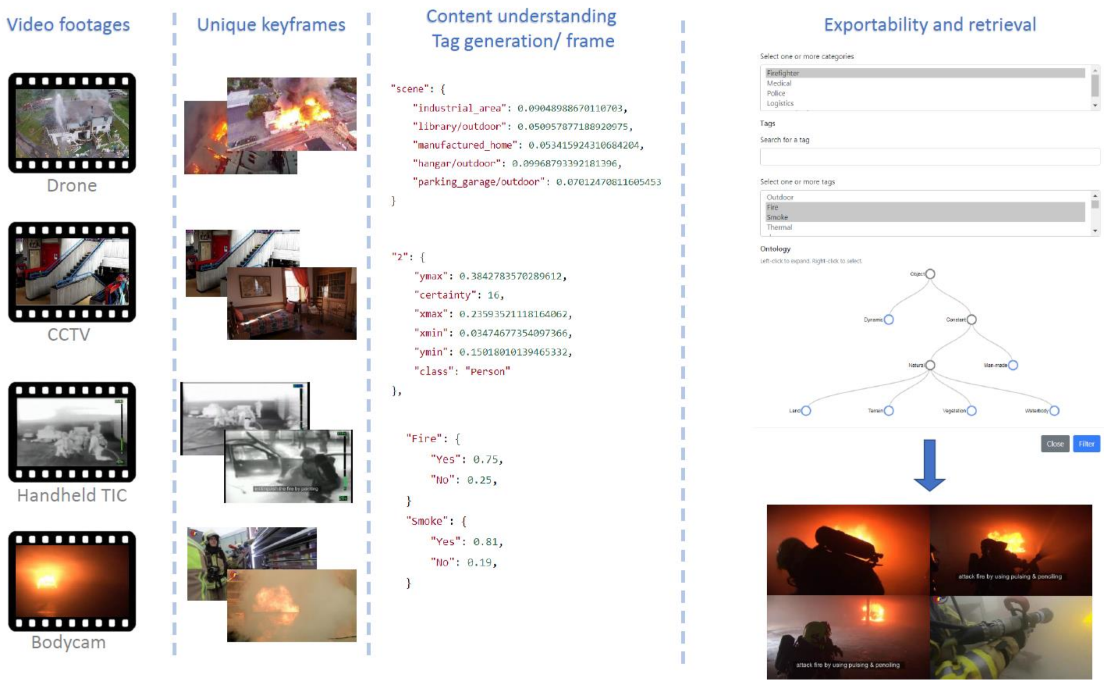

Figure 3. Schematic overview of the video footage analysis framework, first the footage linking, secondly the keyframe generation, thirdly the content understanding tool and finally the frame retrieval tool.

Some improvements for the tag generation are the classification into categories, for example an additional classifier into medical, police, fire service. Furthermore, it is valuable to add the source of the video footage as a tag (e.g., handheld camera first crew, TIC second crew, drone footage). Finally, as indicated earlier in this section, the video summarization and retrieval building blocks require further user-driven evaluations. Still the feasibility of the proposed building blocks is shown in this paper. Real-fire experiments that are recorded from different point-of-views and with different resources are highly valuable for further evaluation and improvement.

6 CONCLUSION

This paper presented the user-needs and data restrictions specific for fire incident management. The insights were gained through the analysis of a questionnaire launched in the firefighting community of Belgium. This evaluation revealed that people are visually-oriented and that video footages are great to gain insights into a problem. Still, people can only process 7 image inputs simultaneous and for that reason, the video summarization framework, consisting of shot detection, frame quality and similarity analysis was proposed. Subsequently, in order to facilitate the video search process, the video and frame retrieval mechanism was clarified and semantic tag based querying on an existing ontology map was initiated. Future work will focus on video analytics to perform person recognition, crowd analysis(e.g., detecting the amount of people and their behavior) and anomaly detection on global scale. Maybe in the future, even before a fire incident will occur the video analysis will notify suspicious behavior and faster interventions could reduce the economic and material damage. Subsequently, more user-evaluations and usability testing will be necessary to make a complete product out of our proposed framework. Therefore the next step is to organize an usability evaluation using the System Usability Scale (SUS)[34]. Finally a closed field test (limited set of users) and a open field test will give insights in the required development steps.

7 ACKNOWLEDGMENTS

The research activities as described in this paper were funded by Ghent University through GOA project BOF16/GOA/004.

8 REFERENCES

[1] M.M. Chun, Contextual cueing of visual attention, Trends in cognitive sciences, 4 (2000) 170-178.
[2] O. Janssens, Data-driven performance monitoring, fault detection and dynamic dashboards for offshore wind farms, in, Ghent University, 2017.
[3] F. Ungerer, H.-J. Schmid, An introduction to cognitive linguistics, Routledge, 2013.
[4] D. Dörner, H. Schaub, Errors in planning and decision-making and the nature of human information processing, Applied psychology, 43 (1994) 433-453.
[5] T. Beji, S. Verstockt, R. Van de Walle, B. Merci, On the use of real-time video to forecast fire growth in enclosures, Fire Technology, 50 (2014) 1021-1040.
[6] F. Vandecasteele, B. Merci, S. Verstockt, Reasoning on multi-sensor geographic smoke spread data for fire development and risk analysis, Fire Safety Journal, 86 (2016) 65-74.
[7] N. Li, Z. Yang, A. Ghahramani, B. Becerik-Gerber, L. Soibelman, Situational awareness for supporting building fire emergency response: Information needs, information sources, and implementation requirements, Fire safety journal, 63 (2014) 17-28.
[8] V. Nunavath, A. Prinz, T. Comes, J. Radianti, Representing fire emergency response knowledge through a domain modelling approach, (2016).
[9] F. Vandecasteele, K. Vandenbroucke, D. Schuurman, S. Verstockt, Spott: On-the-Spot e-Commerce for Television Using Deep Learning-Based Video Analysis Techniques, ACM Trans. Multimedia Comput. Commun. Appl., 13 (2017) 1-16.
[10] F. Vandecasteele, B. Merci, S. Verstockt, Fireground location understanding by semantic linking of visual objects and building information models, Fire Safety Journal, 91 (2017) 1026-1034.
[11] M. Ajmal, M.H. Ashraf, M. Shakir, Y. Abbas, F.A. Shah, Video summarization: techniques and classification, in: International Conference on Computer Vision and Graphics, Springer, 2012, pp. 1-13.
[12] F. Wang, C.-W. Ngo, Summarizing rushes videos by motion, object, and event understanding, IEEE Transactions on Multimedia, 14 (2012) 76-87.
[13] L. dos Santos Belo, C.A. Caetano Jr, Z.K.G. do Patrocínio Jr, S.J.F. Guimarães, Summarizing video sequence using a graph-based hierarchical approach, Neurocomputing, 173 (2016) 1001-1016.
[14] L. Baraldi, C. Grana, R. Cucchiara, Shot and scene detection via hierarchical clustering for re-using broadcast video, in: International Conference on Computer Analysis of Images and Patterns, Springer, 2015, pp. 801-811.
[15] X. Qiu, S. Jiang, H. Liu, Q. Huang, L. Cao, Spatial-temporal attention analysis for home video, in: 2008 IEEE International Conference on Multimedia and Expo (ICME), IEEE, 2008, pp. 1517-1520.
[16] M.R. Mishra, S. Singhai, M. Sharma, A Comparative based study of Different Video-Shot Boundary Detection algorithms, International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), 2 (2013) pp: 282-289.
[17] K. Joy, E.G. Sarma, RECENT DEVELOPMENTS IN IMAGE QUALITY ASSESSMENT ALGORITHMS: A REVIEW, Journal of Theoretical & Applied Information Technology, 65 (2014).

[18] S. Zagoruyko, N. Komodakis, Learning to compare image patches via convolutional neural networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 4353-4361.
[19] S. Verstockt, S. Nop, F. Vandecasteele, T. Baert, N. Van de Weghe, H. Paulussen, E. Rizza, M. Roeges, UGESCO-A Hybrid Platform for Geo-Temporal Enrichment of Digital Photo Collections Based on Computational and Crowdsourced Metadata Generation, in: Euro-Mediterranean Conference, Springer, 2018, pp. 113-124.
[20] J. Luo, A. Savakis, Indoor vs outdoor classification of consumer photographs using low-level and semantic features, in: Image Processing, 2001. Proceedings. 2001 International Conference on, IEEE, 2001, pp. 745-748.
[21] B. Zhou, A. Lapedriza, J. Xiao, A. Torralba, A. Oliva, Learning deep features for scene recognition using places database, in: Advances in neural information processing systems, 2014, pp. 487-495.
[22] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770-778.
[23] R. Girshick, Fast r-cnn, in: Proceedings of the IEEE international conference on computer vision, 2015, pp. 1440-1448.
[24] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, A.C. Berg, Ssd: Single shot multibox detector, in: European conference on computer vision, Springer, 2016, pp. 21-37.
[25] F. Vandecasteele, B. Merci, A. Jalalvand, S. Verstockt, Object localization in handheld thermal images for fireground understanding, in: Thermosense: Thermal Infrared Applications XXXIX, International Society for Optics and Photonics, 2017, pp. 1021405.
[26] A.E. Çetin, K. Dimitropoulos, B. Gouverneur, N. Grammalidis, O. Günay, Y.H. Habiboğlu, B.U. Töreyin, S. Verstockt, Video fire detection–review, Digital Signal Processing, 23 (2013) 1827-1843.
[27] S. Frizzi, R. Kaabi, M. Bouchouicha, J.-M. Ginoux, E. Moreau, F. Fnaiech, Convolutional neural network for video fire and smoke detection, in: Industrial Electronics Society, IECON 2016-42nd Annual Conference of the IEEE, IEEE, 2016, pp. 877-882.
[28] T. Pedersen, S. Patwardhan, J. Michelizzi, WordNet:: Similarity: measuring the relatedness of concepts, in: Demonstration papers at HLT-NAACL 2004, Association for Computational Linguistics, 2004, pp. 38-41.
[29] F. Zhang, D. Zhang, Y. Liu, H. Lin, Representing place locales using scene elements, Computers, Environment and Urban Systems, (2018).
[30] I. Kadar, O. Ben-Shahar, SceneNet: A perceptual ontology for scene understanding, in: European Conference on Computer Vision, Springer, 2014, pp. 385-400.
[31] J.P.A. Almeida, P.D. Costa, G. Guizzardi, Towards an Ontology of Scenes and Situations, in: 2018 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA), IEEE, 2018, pp. 29-35.
[32] G. Poveda, E. Serrano, M. Garijo, Designing emergency management services by ontology driven social simulation, IT CoNvergence PRActice, 3 (2015) 17-32.
[33] A.S. Murugan, K.S. Devi, A. Sivaranjani, P. Srinivasan, A study on various methods used for video summarization and moving object detection for video surveillance applications, Multimedia Tools and Applications, (2018) 1-18.
[34] J. Brooke, SUS-A quick and dirty usability scale, Usability evaluation in industry, 189 (1996) 4-7.