Radiology Objects in COntext (ROCO): A Multimodal Image Dataset

Obioma Pelka[1,2]
[1]Faculty of Medicine
University of Duisburg-Essen, Germany
[2]Department of Computer Sciences
University of Applied Sciences and Arts Dortmund, Germany
obioma.pelka@fh-dortmund.de

Sven Koitka sven.koitka@uk-essen.de
Johannes Rückert johannes.rueckert@fh-dortmund.de
Felix Nensa felix.nensa@uk-essen.de
Christoph M. Friedrich christoph-friedrich@fh-dortmund.de

1 Introduction

Large-scale and easily accessible datasets are needed to create and evaluate model algorithms for research topics. ImageNET [7], a popular and often applied dataset for image classification, and Microsoft Common Objects in Context (MSCOCO) [5] have helped finding solution to the image classification task. We present in this work a new, free and accessible dataset that concentrates on the detection of contextual object interplay solely in radiology images. The Radiology Objects in Context (ROCO) dataset has two classes: Radiology and Out-Of-Class. The first contains 81,825 radiology images with several medical imaging modalities including, Computer Tomography (CT), Ultrasound, X-Ray, Fluoroscopy, Positron Emission Tomography (PET), Mammography, Magnetic Resonance Imaging (MRI), Angiography and PET-CT. The latter contains 6,127 out-of-class samples, including synthetic radiology figures, digital art and portraits. The corresponding captions, keywords, UMLS (Unified Medical Language System) [1] Semantic Types (SemTypes), UMLS Concept Unique Identifiers (CUIs) and download link is distributed for each image.

2 Material and Methods

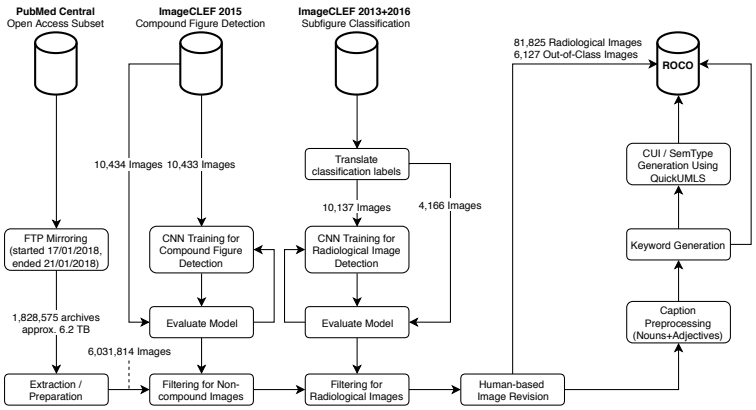

Figure 1: Overview of the work-flow used for the ROCO dataset development [6].

32nd Conference on Neural Information Processing Systems (NIPS 2018), Montréal, Canada.

From the PubMed Open Access subset, a total number of 6,031,814 image - caption pairs were extracted. These images were filtered for non-compound and radiology images using deep learning systems trained with the ImageCLEF 2013, 2015 and 2016 Medical Classification datasets [2, 3, 4]. A final revision was done to manually detect false positives, which include synthetic radiology images, illustrations, portraits, digital artwork, compound radiology images, and make up the out-of-class set.

Semantic knowledge of object interplay present in the images were extracted in form of UMLS Semantic types and Concept by applying QuickUMLS [8]. In addition, the captions were reduced to only nouns and adjectives, which is distributed as 'Keywords'.

3 Results

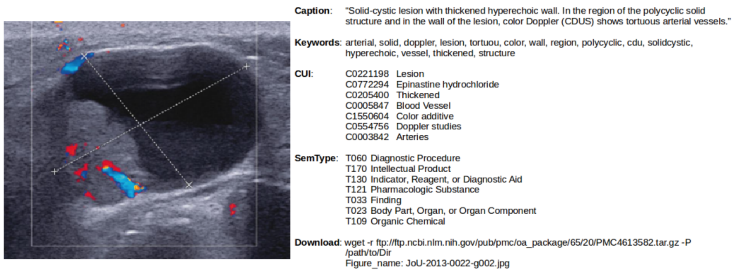

Figure 2: An example of a ROCO image, showing corresponding caption, keywords, UMLS CUIs, UMLS Semantic Types and PMCID download link.

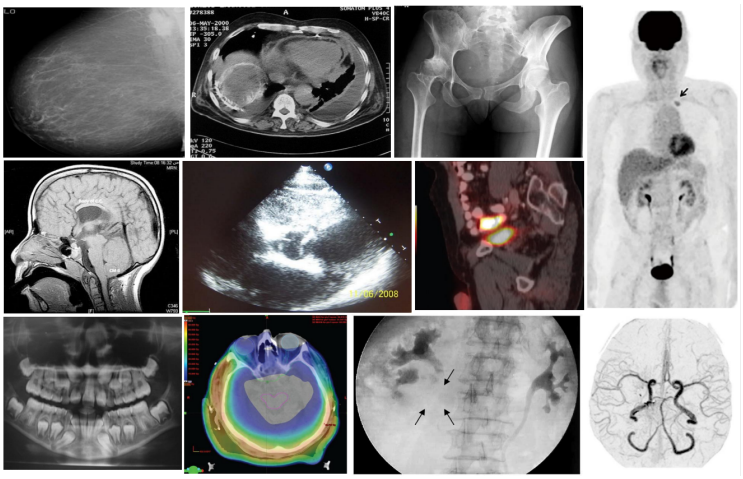

Figure 3: Examples of images contained in the ROCO dataset, illustrating the variety of medical imaging modalities. All images belong to the 'Radiology' subset.

4 Conclusion

The keywords can be adopted as textual features for multi-modal image representations in classification tasks, as well as for multi-class image classification and labeling. Automatic keyword generation models can be designed using ROCO image - keyword pairs, enabling perceivable order for unstructured and unlabeled radiology images and for datasets lacking textual representations. Natural sentences describing radiology images can be created using generative models trained with ROCO image - caption pairs. This will offer additional knowledge of the images and not be limited to solely visual representations.

References

[1] Olivier Bodenreider. The Unified Medical Language System (UMLS): integrating biomedical terminology. *Nucleic Acids Research*, 32(Database-Issue):267–270, 2004.

[2] Alba García Seco de Herrera, Jayashree Kalpathy-Cramer, Dina Demner Fushman, Sameer Antani, and Henning Müller. Overview of the ImageCLEF 2013 medical tasks. In *Working Notes of CLEF 2013 - Conference and Labs of the Evaluation forum, Valencia, Spain, 23-26 September, 2013. CEUR-WS Proceedings Notes, Volume 1179*.

[3] Alba García Seco de Herrera, Henning Müller, and Stefano Bromuri. Overview of the ImageCLEF 2015 medical classification task. In *Working Notes of CLEF 2015 - Conference and Labs of the Evaluation Forum, Toulouse, France, September 8-11, 2015. CEUR-WS Proceedings Notes, Volume 1391*, September 2015.

[4] Alba García Seco de Herrera, Roger Schaer, Stefano Bromuri, and Henning Müller. Overview of the ImageCLEF 2016 medical task. In *Working Notes of CLEF 2016 - Conference and Labs of the Evaluation forum, Évora, Portugal, 5-8 September, 2016. CEUR-WS Proceedings Notes, Volume 1609*, September 2016.

[5] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Lawrence Zitnick. Microsoft COCO: Common Objects in Context. In David Fleet, Tomas Pajdla, Bernt Schiele, and Tinne Tuytelaars, editors, *Proceedings of Computer Vision – ECCV 2014, 13th European Conference, Zurich, Switzerland, September 6-12, 2014*, pages 740–755, Cham, 2014. Springer International Publishing.

[6] Pelka Obioma, Sven Koitka, Johannes Rückert, Felix Nensa, and Christoph M. Friedrich. Radiology objects in context (roco): A multimodal image dataset. In *Large-scale Annotation of Biomedical Data and Expert Label Synthesis (LABELS) – [MICCAI 2018 Satellite Workshop], September 11-14, 2017, Lecture Notes in Computer Science (LNCS)*. Springer International Publishing.

[7] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3):211–252, Dec 2015.

[8] Luca Soldaini and Nazli Goharian. QuickUMLS: a fast, unsupervised approach for medical concept extraction. In *MedIR Workshop, SIGIR*, 2016.