Medical Image Retrieval: Applications and Resources

Henning Müller
University of Applied Sciences Western Switzerland (HES-SO)
Sierre, Switzerland
henning.mueller@hevs.ch

ABSTRACT

Motivation: Medical imaging is one of the largest data producers in the world and over the last 30 years this production increased exponentially via a larger number of images and a higher resolution, plus totally new types of images. Most images are used only in the context of a single patient and a single time point, besides a few images that are used for publications or in teaching. Data are usually scattered across many institutions and cannot be combined even for the treatment of a single patient. Much knowledge is stored in these medical archives of images and other clinical information and content-based medical image retrieval has from the start aimed at making such knowledge accessible using visual information in combination with text or structured data. With the digitization of radiology that started in the mid 1990s the foundation for broader use was laid out.

Problem statement: This keynote presentation aims at giving a historical perspective of how medical image retrieval has evolved from a few prototypes using first only text, then global visual features to the current multimodal systems that can index many types of images in large quantities and use deep learning as a basis for the tools [1,2,3,4]. It also aims at looking at what the place of image retrieval is in medicine, where it is currently still only sparsely used in clinical practice. It seems that it is mainly a tool for teaching and research. Certified medical tools for decision support rather make use of specific approaches for detection and classification.

Approach: The presentation follows a systematic review of the domain that includes many examples of systems and approaches that changed over time when better performing tools became available. Medical mage retrieval has evolved strongly, and many tools linked to mage retrieval are now employed as clinical decision support but mainly for detection and classification. Retrieval remains useful but is often integrated with tools and thus has become almost invisible.

A second aspect of the presentation includes a presentations of existing data sets and other resources that were difficult to obtain even ten years ago, but that have been shared via repositories such as TCGA (The Cancer Genome Atlas, https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga), TCIA (The Cancer Imaging Archive, https://www.cancerimagingarchive.net), or via scientific challenges such ImageCLEF [5] or listed in the Grand Challenges web page (https://grand-challenge.org). Medical data are now easily accessible in many fields and often even in large quantities.

Discussion: Medical retrieval has gone from single text or image retrieval to multimodal approaches [6], really aiming to use all data available for a case, similar to what a physician would do by looking at a patient holistically. The limiting factor in terms of data access is now rather linked to limited manual annotations, as the time of clinicians for annotations is expensive. Global labels for images usually exist with the associated text reports that describe images and outcomes. Still, these weak labels need to be made usable with deep learning approaches that possibly require large amounts of data to generalize well.

Conclusions: Medical image retrieval has evolved strongly over the past 30 years and can be integrated with several tools. For real clinical decision support, it is still rarely used, also because the certification process is tedious and commercial benefit is not as easy to show, as with detection or classification in a clear and limited scenario. In terms of research many resources are available that allow advances also in the future. Still, certification and ethical aspects also need to be taken into account to limit risks for individuals.

CCS Concepts/ACM Classifiers

Information systems, Information retrieval

Author Keywords

Medical image retrieval; visual information retrieval; ImageCLEF; retrieval benchmarking

BIOGRAPHY

Henning Müller studied medical informatics at the University of Heidelberg, Germany, then worked at Daimler-Benz research in Portland, OR, USA. From 1998-2002 he worked on his PhD degree in computer vision (content-based image retrieval) at the University of Geneva, Switzerland with a research stay at Monash University, Melbourne, Australia. Since 2002, Henning has been

ICMR '20, June 8–11, 2020, Dublin, Ireland.

ACM ISBN 978-1-4503-7087-5/20/06.
https://doi.org/10.1145/3372278.3390668

working for the medical informatics service at the University Hospital of Geneva. Since 2007, he has been a full professor at the HES-SO Valais and since 2011 he is responsible for the eHealth unit of the school. Since 2014, he is also professor at the medical faculty of the University of Geneva. In 2015/2016 he was on sabbatical at the Martinos Center, part of Harvard Medical School in Boston, MA, USA to focus on research activities. Henning is coordinator of the ExaMode EU project, was coordinator of the Khresmoi EU project, scientific coordinator of the VISCERAL EU project and is initiator of the ImageCLEF benchmark that has run medical tasks every year since 2004. He has authored over 600 scientific papers with more than 15,000 citations and is in the editorial board of several journals.

REFERENCES

[1] Tang, L. H., Hanka, R., & Ip, H. H. (1999). A review of intelligent content-based indexing and browsing of medical images. *Health Informatics Journal*, *5*(1), 40-49.

[2] Lehmann, T. M., Güld, M. O., Thies, C., Fischer, B., Spitzer, K., Keysers, D., & Wein, B. B. (2004). Content-based image retrieval in medical applications. *Methods of information in medicine*, *43*(04), 354-361.

[3] Müller, H., Michoux, N., Bandon, D., & Geissbuhler, A. (2004). A review of content-based image retrieval systems in medical applications—clinical benefits and future directions. *International journal of medical informatics*, *73*(1), 1-23.

[4] Akgül, C. B., Rubin, D. L., Napel, S., Beaulieu, C. F., Greenspan, H., & Acar, B. (2011). Content-based image retrieval in radiology: current status and future directions. *Journal of digital imaging*, *24*(2), 208-222.

[5] Kalpathy-Cramer, J., de Herrera, A. G. S., Demner-Fushman, D., Antani, S., Bedrick, S., & Müller, H. (2015). Evaluating performance of biomedical image retrieval systems—An overview of the medical image retrieval task at ImageCLEF 2004–2013. *Computerized Medical Imaging and Graphics*, *39*, 55-61.

[6] Müller, H., and Unay, D. (2017). Retrieval from and understanding of large-scale multi-modal medical datasets: a review." *IEEE transactions on multimedia* 19.9: 2093-2104.