Medical Multimedia Systems and Applications

Pål Halvorsen*
SimulaMet, Norway
paalh@simula.no

Michael Alexander Riegler†
SimulaMet, Norway
michael@simula.no

Klaus Schoeffmann
Klagenfurt University, Austria
ks@itec.aau.at

ABSTRACT

In recent years, we have observed a rise of interest in the multimedia community towards research topics related to health. It can be observed that this goes into two interesting directions. One is personal health with a larger focus on well-being and everyday healthy living. The other direction focuses more on multimedia challenges within the health-care systems, for example, how can multimedia content produced in hospitals be used efficiently but also on the user perspective of patients and health-care personal. Challenges and requirements in this interesting and challenging direction are similar to classic multimedia research, but with some additional pitfalls and challenges. This tutorial aims to give a general introduction to the research area; to provide an overview of specific requirements, pitfalls and challenges; to discuss existing and possible future work; and to elaborate on how machine learning approaches can help in multimedia-related challenges to improve the health-care quality for patients and support medical experts in their daily work.

ACM Reference Format:
Pål Halvorsen, Michael Alexander Riegler, and Klaus Schoeffmann. 2019. Medical Multimedia Systems and Applications. In *Proceedings of the 27th ACM International Conference on Multimedia (MM '19), October 21–25, 2019, Nice, France.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3343031.3351319

1 MOTIVATION

The health-care systems produce large amounts of data every day. This ranges from sensor data, over patients records to images and videos. Most of the time, medical experts and patients are not able to make efficient use of all the data collected. Algorithms exist that perform well in detecting for example skin cancer, but most of them focus on single modalities and do not take the users (patients and doctors) or use-cases (how would the output of the algorithm actually be used efficiently?) into account. Specifically, images and videos have gained increasing importance over the last decade. One of the main reasons is the breakthrough of endoscopic imaging in almost all medical specialties in combination with the availability of cheap storage capacity. This encourages more and more surgeons to record their work and use the corresponding *medical videos* for various purposes [6], such as documentation, post-hoc reviewing, analysis of technical errors [1, 2], demonstration of novel operation techniques (in training and teaching), and preparation for follow-up procedures. Also, in a diagnostic context, videos and images are extremely helpful sources of information, for example in gastroscopy and colonoscopy [12]. However, advanced multimedia technology is needed to unfold the full potential of this data. We need effective tools for automatic diagnostic decision support, real-time support during surgeries as well as support in semantic structuring of documentation archives, e.g., by automated understanding of specific medical semantics, like anatomy, surgical actions, conditions, pathologies, etc., by using machine learning [3, 9].

Recently, several medical multimedia datasets have been released [4, 10, 11, 19] and even the newest MPEG video coding standard VVC (Versatile Video Coding) mentions the medical domain as one of its target fields [7]. Both examples are emphasizing the importance of efficient analysis and management of multimedia data in the area of medicine.

2 TUTORIAL

This is the third installment [13, 18] of a tutorial focusing on the field of medical multimedia. In this tutorial we present existing methods to approach several medical challenges [14]. This includes a detailed description of the data and its specifics characteristics, content analysis methods proposed for real-time detection as well as post-processing for content structuring, semantic classification, and information retrieval.

The rapidly increasing amount of data that is collected and the growing interest of health professionals necessitate the conception and development of specifically domain-adapted interactive multimedia systems that provide support for diagnostics, examination, surgery, reporting and teaching in a medical setting by combining all available information sources and putting them in the hands of medical professionals or patients. As such, a complete system has to fulfill several requirements, which include efficient processing capabilities, a pipeline for the complete workflow from medical expert knowledge transfer using annotation tools via an automatic analysis module through to a component for presentation of the systems' output and the possibility to adapt and extend the system regarding different requirements such as different diseases. The tutorial also covers available datasets and latest achievements in learning semantics from the medical domain going beyond single modalities. Content analysis of medical data is a very challenging area, since these videos and images often contain a very special content. For example, the actual content area is often circular (the frame is surrounded by a black border) and contains a considerable portion of irrelevant content. More importantly, such videos contain highly self-similar content (similar colors, similar texture, similar motion) due to highly similar anatomical structures as well

*Also affiliated with Oslo Metropolitan University, Norway
†Also affiliated with Kristiania University College, Norway

MM '19, October 21–25, 2019, Nice, France

ACM ISBN 978-1-4503-6889-6/19/10.
https://doi.org/10.1145/3343031.3351319

as similar actions and operations. Therefore, it is very challenging to create discriminative content features and sufficiently large and diverse training sets when using machine learning techniques. Furthermore, special interfaces for interaction with medical multimedia data are presented. In this context, it is important to note that doctors also need an understanding of the automated decision support, and the system may also assist the administrative tasks after the examinations.

Participants are provided with best practices on how to process medical multimedia content, what tools and features clinicians need and what they expect, and how multimedia research can help. They should be able to apply this knowledge in their own projects upon completion of the tutorial.

2.1 Duration and Modules

The tutorial is conducted as half-day tutorial, contains several medical examples and covers (at least) topics as follows:

(1) Introduction and Motivation
   (a) What and Why: Medical Multimedia
   (b) Characteristics, Challenges & Pitfalls
   (c) Use Cases and Application Domains
   (d) Research Areas of Medical Multimedia

(2) State-of-the-Art in Medical Multimedia [6]
   (a) Pre-processing Methods
   (b) Machine Learning Methods
   (c) Real-time applications
   (d) Post-Procedural applications

(3) Relevant Medical Datasets [4, 10, 11, 19]
   (a) Challenges and Requirements
   (b) Ethical Issues
   (c) Public Datasets: How to?

(4) Diagnostic Decision Support [5, 14, 15]
   (a) Knowledge Transfer from Medical Experts
   (b) Real-Time Feedback
   (c) Challenges and Best Practices

(5) Medical Multimedia Applications [8, 16, 17]
   (a) Smart Documentation
   (b) Concept and Scene Understanding
   (c) Surgical Quality Assessment
   (d) Interaction, Browsing & Exploration

(6) Conclusions and Future Challenges

2.2 Target Audience, Attendees, Material

Our target audience is researchers and practitioners working in at least one of the following fields: video content analysis, visual information retrieval, machine learning for multimedia, and human-computer interaction with multimedia data. Participants should have a basic understanding of multimedia content analysis and should be somewhat familiar with machine learning methods. We do not expect any medical background. One of our aims is to enthuse participants with this new application domain for multimedia research and stimulate further research in this area.

3 SUMMARY

This tutorial provides a broad overview of the area of medical multimedia. This includes the medical social challenges, the existing medical data, the potential of the large knowledge-base in the field of multimedia, existing examples of medical-multimedia interdisciplinary work with existing methods and solutions and the open challenges that fits perfectly to the multimedia community. We hope to raise awareness of the huge potential and growing significance of medical multimedia research and encourage further cross-domain research between the communities – potentially enabling multimedia researchers to contribute to save human lives in the future.

4 PRESENTERS

Michael Alexander Riegler is a senior researcher at SimulaMet and an adjunct associate Professor at the Kristiania University College. He received his Master's degree from Klagenfurt University with distinction and finished his PhD at the University of Oslo in two and a half years. His PhD thesis topic was efficient processing of medical multimedia workloads. His research interests are medical multimedia data analysis and understanding, image processing, image retrieval, parallel processing, crowdsourcing, social computing, and user intent. He is involved in several initiatives like the MediaEval Benchmarking Initiative for Multimedia Evaluation, which runs the Medico task. Furthermore, he is part of an expert group for the Norwegian Council of Technology on Machine Learning for health-care.

Pål Halvorsen is a chief research scientist at SimulaMet, professor at Oslo Metropolitan University, professor II at University of Oslo, and the CEO of Forzasys AS (all Norway). He received his doctoral degree in 2001. His research focuses mainly at distributed multimedia systems including operating systems, processing, storage and retrieval, communication and distribution from a performance and efficiency point of view. He has co-authored more than 200 publications on various topics in multimedia. More information about authored papers, projects, supervised students, teaching, community services, etc. can be found at http://home.ifi.uio.no/paalh.

Klaus Schoeffmann is an associate professor at the Institute of Information Technology (ITEC) at Klagenfurt University, Austria, where he received his habilitation (venia docendi) in Computer Science in 2015. He holds a PhD (with distinction) and a MSc (with distinction) in Computer Science. His research focuses on video content analysis (in particular of medical/surgery videos), multimedia retrieval, interactive multimedia, and deep learning. He has co-authored more than 120 publications on various topics in multimedia, inclusive of more than 30 on different aspects of medical video analysis. He has co-organized several international conferences, workshops, and special sessions in the field of multimedia.

REFERENCES

[1] EM Bonrath, B Zevin, NJ Dedy, and TP Grantcharov. 2013. Error rating tool to identify and analyse technical errors and events in laparoscopic surgery. *British Journal of Surgery* 100, 8 (2013), 1080–1088.

[2] Nicolas J Dedy, Peter Szasz, Marisa Louridas, Esther M Bonrath, Heinrich Husslein, and Teodor P Grantcharov. 2015. Objective structured assessment of non-technical skills: reliability of a global rating scale for the in-training assessment in the operating room. *Surgery* 157, 6 (2015), 1002–1013.

[3] Sabrina Kletz, Klaus Schoeffmann, Jenny Benois-Pineau, and Heinrich Husslein. 2019. Identifying Surgical Instruments in Laparoscopy Using Deep Learning Instance Segmentation. In *2019 International Conference on Content-Based Multimedia Indexing (CBMI)*. IEEE, 1–6.

[4] Andreas Leibetseder, Stefan Petscharnig, Manfred Jürgen Primus, Sabrina Kletz, Bernd Münzer, Klaus Schoeffmann, and Jörg Keckstein. 2018. LapGyn4: A Dataset for 4 Automatic Content Analysis Problems in the Domain of Laparoscopic Gynecology. In *Proceedings of the ACM International Conference on Multimedia Retrieval (ICMR)*. ACM, 1–6. to appear.

[5] Andreas Leibetseder, Manfred Jürgen Primus, Stefan Petscharnig, and Klaus Schoeffmann. 2017. Image-Based Smoke Detection in Laparoscopic Videos. In *Computer Assisted and Robotic Endoscopy and Clinical Image-Based Procedures*, M. Jorge Cardoso, Tal Arbel, Xiongbiao Luo, Stefan Wesarg, Tobias Reichl, Miguel Ángel González Ballester, Jonathan McLeod, Klaus Drechsler, Terry Peters, Marius Erdt, Kensaku Mori, Marius George Linguraru, Andreas Uhl, Cristina Oyarzun Laura, and Raj Shekhar (Eds.). Springer International Publishing, Cham, 70–87.

[6] Bernd Münzer, Klaus Schoeffmann, and Laszlo Böszörmenyi. 2018. Content-based processing and analysis of endoscopic images and videos: A survey. *Multimedia Tools and Applications* 77, 1 (2018), 1323–1362.

[7] Jens-Rainer Ohm and Gary J Sullivan. 2018. Versatile Video Coding–towards the next generation of video compression. In *Picture Coding Symposium 2018*.

[8] Stefan Petscharnig and Klaus Schöffmann. 2017. Learning laparoscopic video shot classification for gynecological surgery. *Multimedia Tools and Applications* (2017), 1–19.

[9] Stefan Petscharnig and Klaus Schöffmann. 2018. Learning laparoscopic video shot classification for gynecological surgery. *Multimedia Tools and Applications* 77, 7 (01 Apr 2018), 8061–8079. https://doi.org/10.1007/s11042-017-4699-5

[10] Konstantin Pogorelov, Kristin Ranheim Randel, Thomas de Lange, Sigrun Losada Eskeland, Carsten Griwodz, Dag Johansen, Concetto Spampinato, Mario Taschwer, Mathias Lux, Peter Thelin Schmidt, Michael Riegler, and Pål Halvorsen. 2017. Nerthus: A Bowel Preparation Quality Video Dataset. In *Proceedings of ACM Multimedia Systems (MMSYS)*.

[11] Konstantin Pogorelov, Kristin Ranheim Randel, Carsten Griwodz, Sigrun Losada Eskeland, Thomas de Lange, Dag Johansen, Concetto Spampinato, Duc-Tien Dang-Nguyen, Mathias Lux, Peter Thelin Schmidt, Michael Riegler, and Pål Halvorsen. 2017. Kvasir: A Multi-Class Image Dataset for Computer Aided Gastrointestinal Disease Detection. In *Proceedings of ACM Multimedia Systems (MMSYS)*.

[12] Konstantin Pogorelov, Michael Riegler, Pål Halvorsen, Thomas de Lange, Peter Thelin Schmidt, Sigrun L. Eskeland, Carsten Griwodz, and Dag Johansen. 2016. GPU-accelerated Real-time Gastrointestinal Diseases Detection. In *Proceedings of International Symposium on Computer-Based Medical Systems (CBMS)*. IEEE.

[13] Michael Riegler, Pål Halvorsen, Bernd Münzer, and Klaus Schoeffmann. 2018. The Importance of Medical Multimedia. In *Proceedings of the 26th ACM International Conference on Multimedia (MM '18)*. ACM, New York, NY, USA, 2106–2108. https://doi.org/10.1145/3240508.3241475

[14] Michael Riegler, Mathias Lux, Carsten Griwodz, Concetto Spampinato, Thomas de Lange, Sigrun L. Eskeland, Konstantin Pogorelov, Wallapak Tavanapong, Peter T. Schmidt, Cathal Gurrin, Dag Johansen, Håvard Johansen, and Pål Halvorsen. 2016. Multimedia and Medicine: Teammates for Better Disease Detection and Survival. In *Proceedings of the ACM on Multimedia Conference (ACM MM)*. 968–977. https://doi.org/10.1145/2964284.2976760

[15] Michael Riegler, Konstantin Pogorelov, Sigrun L. Eskeland, Peter T. Schmidt, Zeno Albisser, Dag Johansen, Carsten Griwodz, Pål Halvorsen, and Thomas de Lange. 2017. From Annotation to Computer Aided Diagnosis: Detailed Evaluation of a Medical Multimedia System. *ACM Transactions on Multimedia Computing Communications and Applications* 13, 3 (2017).

[16] Klaus Schoeffmann, Christian Beecks, Mathias Lux, Merih Seran Uysal, and Thomas Seidl. 2016. Content-based retrieval in videos from laparoscopic surgery. In *SPIE Medical Imaging*. International Society for Optics and Photonics, 97861V–97861V.

[17] Klaus Schoeffmann, Marco A. Hudelist, and Jochen Huber. 2015. Video Interaction Tools: A Survey of Recent Work. *ACM Comput. Surv.* 48, 1, Article 14 (Sept. 2015), 34 pages. https://doi.org/10.1145/2808796

[18] Klaus Schoeffmann, Bernd Münzer, Michael Riegler, and Pål Halvorsen. 2017. Medical Multimedia Information Systems (MMIS). In *Proceedings of the 25th ACM International Conference on Multimedia (MM '17)*. ACM, New York, NY, USA, 1957–1958. https://doi.org/10.1145/3123266.3130142

[19] Klaus Schoeffmann, Mario Taschwer, Stephanie Sarny, Bernd Münzer, Manfred Jürgen Primus, and Doris Putzgruber. 2018. Cataract-101–Video Dataset of 101 Cataract Surgeries. In *Proceedings of the ACM International Conference on Multimedia Retrieval (ICMR)*. ACM, 1–6. to appear.