ImageCLEF 2019: Multimedia Retrieval in Lifelogging, Medical, Nature, and Security Applications

Bogdan Ionescu[1], Henning Müller[2], Renaud Péteri[3], Duc-Tien Dang-Nguyen[16], Luca Piras[5], Michael Riegler[6], Minh-Triet Tran[7], Mathias Lux[8], Cathal Gurrin[4], Yashin Dicente Cid[2], Vitali Liauchuk[9], Vassili Kovalev[9], Asma Ben Abacha[11], Sadid A. Hasan[10], Vivek Datla[10], Joey Liu[10], Dina Demner-Fushman[11], Obioma Pelka[12], Christoph M. Friedrich[12], Jon Chamberlain[13], Adrian Clark[13], Alba García Seco de Herrera[13], Narciso Garcia[14], Ergina Kavallieratou[15], Carlos Roberto del Blanco[14], Carlos Cuevas Rodríguez[14], Nikos Vasillopoulos[15], and Konstantinos Karampidis[15]

[1] University Politehnica of Bucharest, Romania
bionescu@alpha.imag.pub.ro
[2] University of Applied Sciences Western Switzerland (HES-SO), Switzerland
[3] University of La Rochelle, France
[4] Dublin City University, Ireland
[5] University of Cagliari, Italy
[6] University of Oslo, Norway
[7] University of Science, Vietnam
[8] Klagenfurt University, Austria
[9] Institute for Informatics, Belarus
[10] Philips Research Cambridge, USA
[11] National Library of Medicine, USA
[12] University of Applied Sciences and Arts, Germany
[13] University of Essex, UK
[14] E.T.S. Ingenieros Telecomunicación, Spain
[15] University of the Aegean, Greece
[16] University of Bergen, Norway

Abstract. This paper presents an overview of the foreseen ImageCLEF 2019 lab that will be organized as part of the Conference and Labs of the Evaluation Forum - CLEF Labs 2019. ImageCLEF is an ongoing evaluation initiative (started in 2003) that promotes the evaluation of technologies for annotation, indexing and retrieval of visual data with the aim of providing information access to large collections of images in various usage scenarios and domains. In 2019, the 17th edition of Image-CLEF will run four main tasks: (i) a *Lifelog* task (videos, images and other sources) about daily activities understanding, retrieval and summarization, (ii) a *Medical* task that groups three previous tasks (caption analysis, tuberculosis prediction, and medical visual question answering) with newer data, (iii) a new *Coral* task about segmenting and labeling collections of coral images for 3D modeling, and (iv) a new *Security* task addressing the problems of automatically identifying forged content

and retrieve hidden information. The strong participation, with over 100 research groups registering and 31 submitting results for the tasks in 2018 shows an important interest in this benchmarking campaign and we expect the new tasks to attract at least as many researchers for 2019.

Keywords: lifelogging retrieval and summarization · medical retrieval · coral image segmentation and classification · file forgery detection · ImageCLEF benchmarking · annotated datasets

1   Introduction

The ImageCLEF evaluation campaign was started as part of the CLEF (Cross Language Evaluation Forum) in 2003 [4, 5]. It has been held every year since then and delivered many results in the analysis and retrieval of images [17, 15]. Medical tasks started in 2004 and have in some years been the majority of the tasks in ImageCLEF [14].

The objectives of ImageCLEF have always been the multilingual or language-independent analysis of visual content. A focus has often been on multimodal data sets, so combining images with structure information, free text or other information that helps in the decision making.

Since 2018 ImageCLEF uses the crowdAI[17] platform to distribute the data and received the submitted results. The system allows having an online leader board and gives the possibility to keep data sets accessible beyond competition, including a continuous submission to the leader board.

Over the years, ImageCLEF and also CLEF have shown a strong scholarly impact that was captured in [21, 22]. This underlines the importance of evaluation campaigns for disseminating best scientific practices.

In the following, we introduce the four tasks that are going to run in the 2019 edition[18], namely: ImageCLEFlifelog, ImageCLEFmedical, ImageCLEFcoral, and ImageCLEFsecurity. A sample of some of the provided visual data is presented in Figure 1.

2   ImageCLEFlifelog

An increasingly wide range of personal devices, such as smart-phones, video cameras as well as wearable devices that allow capturing pictures, videos, and audio clips for every moment of our lives have become available. Considering the huge volume of data created, there is a need for systems that can automatically analyze the data in order to categorize, summarize and also query to retrieve the information the user may need.

The main goal of the Lifelog task since its first edition [6] has been to advance the state-of-the-art research in lifelogging as an application of information

---

[17] http://www.crowdA.org/
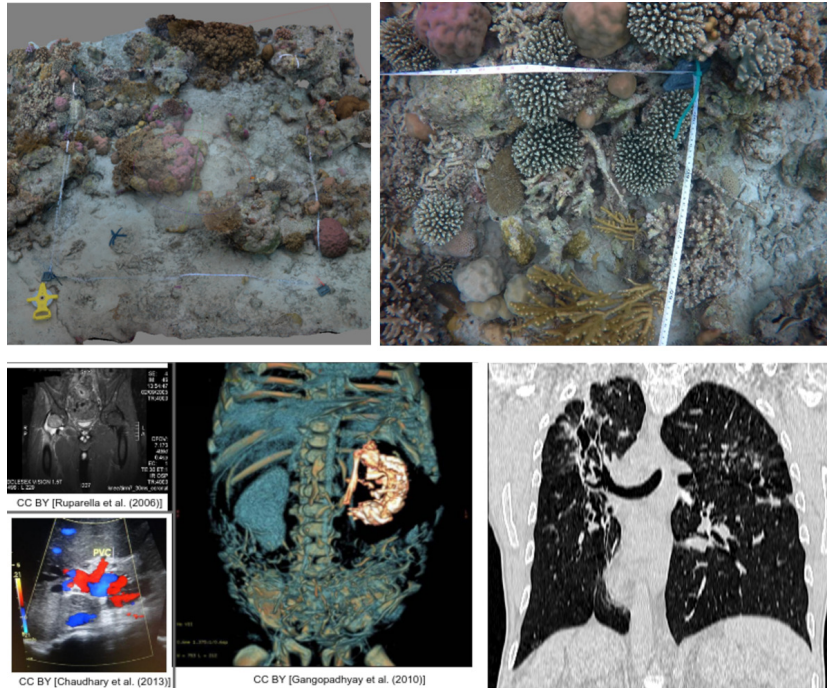[18] https://www.imageclef.org/2019

Fig. 1. Sample images from (left to right, top to bottom): ImageCLEFcoral, and ImageCLEFmedical, caption and tuberculosis tasks.

retrieval. As in the 2017 and 2018 editions, the 2019 task will be split into two related subtasks using a completely new rich multimodal data set. It consists of 42 days of data from two lifeloggers, namely: images (1,500-2,500 per day from wearable cameras), visual concepts (automatically extracted visual concepts with varying rates of accuracy), semantic content (semantic locations, semantic activities) based on sensor readings (via the Moves App) on mobile devices, biometrics information (heart rate, galvanic skin response, calories burn, steps, continual blood glucose, etc.), music listening history, computer usage (frequency of typed words via the keyboard and information consumed on the computer via Automatic Speech Recognition of on-screen activity on a per-minute basis). The copyright and ethical approval to release the data is held by one of the task organizers.

Subtask 1 (Puzzle): Solve my life puzzle. Given a set of lifelog images with associated metadata (e.g., biometrics, location, etc.), but no time stamps, the participants need to analyze these images and rearrange them in chronological order and predict the correct day (Monday or Sunday) and part of the day (morning, afternoon, or evening). The data set will be arranged into 75% training and 25% test data.

Subtask 2 (LMRT): Lifelog moment retrieval. This sub-task follows the success of the LMRT sub-task in ImageCLEFlifelog 2018 [7] with some minor adjustments. The participants have to retrieve a number of specific predefined activities in a lifelogger's life. For example, they need to return the relevant moments for the query "Find the moment(s) when I was shopping". Particular attention needs to be paid to the diversification of the selected moments with respect to the target scenario. The ground truth for this subtask was created using manual annotations.

For assessing performance, classic metrics will be deployed, e.g., precision, cluster recall (to account for the diversification), etc. In particular, the organizers would like to emphasize methods that allow interaction with real users (via Relevance Feedback - RF, for example) and the organizers will define appropriate evaluation measures.

3   ImageCLEFmedical

The medical tasks of ImageCLEF have started in 2004 and have been run almost every year since then [15]. In 2019 there will be three subtasks under the medical umbrella that will all be based on past tasks but concentrating on clean data and on bringing people of the three tasks together with fewer actual subtasks. The three tasks will be: *figure caption analysis* [8, 13], *tuberculosis analysis* [8, 13], and *visual question answering* [12].

The *caption analysis task* will use a new and manually curated data set of images from the biomedical literature, thus reducing variability in the data and making the extraction of concepts cleaner, as only clinical images are present and as there are also quality constraints on the captions. The Radiology Objects in Context (ROCO) [18] data set is used. It contains over 81,000 radiology images from the medical literature including caption information and a manual control of the image type. The task will concentrate on extracting Unified Medical Language System (UMLS) concepts and not the prediction of a precise caption. Trivial concepts will be removed and also concepts occurring in only a single image.

The *tuberculosis task* uses 3D image volumes (Computed Tomography with 3 mm slice thickness and around 150 slices per image volume) and clinical data to detect tuberculosis type and severity from these data. The multiple drug resistance task was dropped for 2019, as results were of limited quality. The other two tasks are clinically more interesting.

The *medical Visual Question Answering (VQA) task* is an exciting problem that combines natural language processing and computer vision techniques. Inspired by the recent success of visual question answering in the general domain, this year task will focus on a new, larger and nicely manually curated dataset. Given a medical image accompanied with a clinically relevant question, participating systems are tasked with answering the question based on the visual image content.

4   ImageCLEFcoral

Most coral reefs are in danger of being lost within the next 30 years, and with them the ecosystems they support [1]. This catastrophe will see the extinction of many marine species, such as shellfish, corals and many micro-organisms in the ocean. It also reduces reef fishery production, which is an important source of income and food source [2, 19]. By monitoring the changes in the structural complexity and composition of coral reefs we can help prioritize conservation efforts. Autonomous Underwater Vehicles (AUV) can collect data for many hours at a time. However, the complexity of the images makes it impossible for human annotators to assess the contents of images on a large scale [3]. Advances in automatically annotating images for complexity and benthic composition have been promising [20, 11], and we are interested in automatically identifying areas of interest and to label them appropriately for monitoring coral reefs.

Similar to previous ImageCLEF annotation tasks [10, 9, 24, 25, 23], the 2019 ImageCLEFcoral task will require participants to automatically annotate and localize a collection of images with types of benthic substrate, such as hard coral and sponge. The data for this task originates from a growing, large-scale collection of images taken from coral reefs around the world as part of a coral reef monitoring project with the Marine Technology Research Unit at the University of Essex (currently containing over 2TB of image data of benthic reef structure).

The annotated data set comprises several sets of overlapping images, each set taken in an area of underwater terrain. Each image will be labelled by experts for training and evaluation.

The performance of the algorithms will be evaluated using the PASCAL VOC[19] style metric of intersection over union (IoU) that calculates the area of intersection between the foreground in the proposed output localization and the ground-truth bounding box localization, divided by the area of their union.

5   ImageCLEFsecurity

File Forgery Detection (FFD) is a serious problem concerning digital forensics examiners. Fraud or counterfeits are common causes for altering files. Another example is a child predator who hides porn images by altering the image extension and in some cases by changing the image signature. Many proposals have been made to solve this problem and the most promising ones concentrate on the image content. It is also common that someone who wants to hide information in plain sight without being perceived might use steganography. Steganography is the practice of concealing a file, message, image or video within another file, message, image, or video. The most usual cover medium for hiding data are images. For more information, we refer the reader to [16].

The specific objective of this task is first to examine if an image has been forged and then if it could hide a text message. Last objective is to retrieve the potentially hidden message from the forged steganography images.

---

[19] http://host.robots.ox.ac.uk/pascal/VOC/

The participant takes the role of a professional digital forensic examiner collaborating with the police, who suspects that there is an ongoing fraud in the Central Bank. After obtaining a court order, police gain access to a suspects computer in the bank with the purpose of looking for images proving the suspect guilty. However, police suspects that the suspect managed to change file extensions and signatures of some images, so that they look like PDF (Portable Document Format) files or other types. It is probable that the suspect has used steganography software to hide messages within the forged images that can reveal valuable information. The following subtaks are defined.

Subtask 1: perform detection of altered (forged) images (both extension and signature) and predict the actual type of the forged file. *Subtask 2*: identify the altered images that hide steganographic content. *Subtask 3*: retrieve the hidden messages (text) from the forged steganographic images.

The data set consists of about 9,000 forged images and pdfs, divided into 3 groups of 3,000 images each one. Every group of images is used for a specific task, where 2,000 images are for training and 1,000 for test. All participants have access to the training data set along with the ground truth. The participants will also have the opportunity to publish an extended version of their proposed methodology and experiments in a special issue of the Journal of Imaging.

6   Conclusions

In this paper we presented an overview of the upcoming ImageCLEF 2019 campaign. ImageCLEF has organized many tasks in a variety of domains over the past 17 years, from general stock photography, medical and biodiversity data to multimodal lifelogging. The focus has always been on language independent approaches and most often on multimodal data analysis. 2019 has a set of interesting tasks that are expected to again draw a large number of participants. A focus for 2019 has been on the diversity of applications and on creating clean data sets to provide a solid basis for the evaluations.

References

1. Birkeland, C.: Global status of coral reefs: In combination, disturbances and stressors become ratchets. In: World Seas: an Environmental Evaluation, pp. 35–56. Elsevier (2019)
2. Brander, L.M., Rehdanz, K., Tol, R.S., Van Beukering, P.J.: The economic impact of ocean acidification on coral reefs. Climate Change Economics **3**(01), 1250002 (2012)
3. Bullimore, R.D., Foster, N.L., Howell, K.L.: Coral-characterized benthic assemblages of the deep northeast atlantic: defining coral gardens to support future habitat mapping efforts. ICES Journal of Marine Science **70**(3), 511–522 (2013)
4. Clough, P., Müller, H., Sanderson, M.: The CLEF 2004 cross–language image retrieval track. In: Peters, C., Clough, P., Gonzalo, J., Jones, G.J.F., Kluck, M., Magnini, B. (eds.) Multilingual Information Access for Text, Speech and Images: Result of the fifth CLEF evaluation campaign. Lecture Notes in Computer Science (LNCS), vol. 3491, pp. 597–613. Springer, Bath, UK (2005)

5. Clough, P., Sanderson, M.: The CLEF 2003 cross language image retrieval task. In: Proceedings of the Cross Language Evaluation Forum (CLEF 2003) (2004)
6. Dang-Nguyen, D.T., Piras, L., Riegler, M., Boato, G., Zhou, L., Gurrin, C.: Overview of ImageCLEFlifelog 2017: Lifelog Retrieval and Summarization. In: Cappellato, L., Ferro, N., Goeuriot, L., Mandl, T. (eds.) CLEF 2017 Working Notes. CEUR Workshop Proceedings (CEUR-WS.org), ISSN 1613-0073, http://ceur-ws.org/Vol-1866/ (2017)
7. Dang-Nguyen, D.T., Piras, L., Riegler, M., Zhou, L., Lux, M., Gurrin, C.: Overview of ImageCLEFlifelog 2018: Daily Living Understanding and Lifelog Moment Retrieval. In: Cappellato, L., Ferro, N., Goeuriot, L., Mandl, T. (eds.) CLEF 2018 Working Notes. CEUR Workshop Proceedings (CEUR-WS.org), ISSN 1613-0073, http://ceur-ws.org/Vol-1866/ (2018)
8. Eickhoff, C., Schwall, I., García Seco de Herrera, A., Müller, H.: Overview of Image-CLEFcaption 2017 - the image caption prediction and concept extraction tasks to understand biomedical images. In: CLEF2017 Working Notes. CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, Dublin, Ireland (September 11-14 2017)
9. Gilbert, A., Piras, L., Wang, J., Yan, F., Dellandrea, E., Gaizauskas, R.J., Villegas, M., Mikolajczyk, K.: Overview of the ImageCLEF 2015 scalable image annotation, localization and sentence generation task. In: CLEF Working Notes (2015)
10. Gilbert, A., Piras, L., Wang, J., Yan, F., Ramisa, A., Dellandrea, E., Gaizauskas, R.J., Villegas, M., Mikolajczyk, K.: Overview of the ImageCLEF 2016 scalable concept image annotation task. In: CLEF Working Notes. pp. 254–278 (2016)
11. Gonzalez-Rivero, M., Bongaerts, P., Beijbom, O., Pizarro, O., Friedman, A., Rodriguez-Ramirez, A., Upcroft, B., Laffoley, D., Kline, D., Bailhache, C., Vevers, R., Hoegh-Guldberg, O.: The catlin seaview surveykilometre-scale seascape assessment, and monitoring of coral reef ecosystems. Aquatic Conservation: Marine and Freshwater Ecosystems **24**, 184–198 (11 2014)
12. Hasan, S.A., Ling, Y., Farri, O., Liu, J., Lungren, M., Müller, H.: Overview of the ImageCLEF 2018 medical domain visual question answering task. In: CLEF2018 Working Notes. CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, Avignon, France (September 10-14 2018)
13. García Seco de Herrera, A., Eickhoff, C., Andrearczyk, V., , Müller, H.: Overview of the ImageCLEF 2018 caption prediction tasks. In: CLEF2018 Working Notes. CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, Avignon, France (September 10-14 2018)
14. Ionescu, B., Müller, H., Villegas, M., de Herrera, A.G.S., Eickhoff, C., Andrearczyk, V., Cid, Y.D., Liauchuk, V., Kovalev, V., Hasan, S.A., Ling, Y., Farri, O., Liu, J., Lungren, M., Dang-Nguyen, D.T., Piras, L., Riegler, M., Zhou, L., Lux, M., Gurrin, C.: Overview of ImageCLEF 2018: Challenges, datasets and evaluation. In: Experimental IR Meets Multilinguality, Multimodality, and Interaction. Proceedings of the Ninth International Conference of the CLEF Association (CLEF 2018), LNCS Lecture Notes in Computer Science, Springer, Avignon, France (September 10-14 2018)
15. Kalpathy-Cramer, J., García Seco de Herrera, A., Demner-Fushman, D., Antani, S., Bedrick, S., Müller, H.: Evaluating performance of biomedical image retrieval systems: Overview of the medical image retrieval task at ImageCLEF 2004–2014. Computerized Medical Imaging and Graphics **39**(0), 55 – 61 (2015)
16. Karampidis, K., Kavallieratou, E., Papadourakis, G.: A review of image steganalysis techniques for digital forensics. Journal of Information Security and Applications **40** (05 2018). https://doi.org/10.1016/j.jisa.2018.04.005

17. Müller, H., Clough, P., Deselaers, T., Caputo, B. (eds.): ImageCLEF – Experimental Evaluation in Visual Information Retrieval, The Springer International Series On Information Retrieval, vol. 32. Springer, Berlin Heidelberg (2010)
18. Pelka, O., Koitka, S., Rückert, J., Nensa, F., Friedrich, C.M.: Radiology Objects in COntext (ROCO): a multimodal image dataset. In: Proceedings of the Third International Workshop on Large-Scale Annotation of Biomedical Data and Expert Label Synthesis (LABELS 2018), Held in Conjunction with MICCAI 2018. vol. 11043, pp. 180–189. LNCS Lecture Notes in Computer Science, Springer, Granada, Spain (September 16 2018)
19. Speers, A.E., Besedin, E.Y., Palardy, J.E., Moore, C.: Impacts of climate change and ocean acidification on coral reef fisheries: An integrated ecological–economic model. Ecological economics **128**, 33–43 (2016)
20. Stokes, M., Deane, G.: Automated processing of coral reef benthic images. Limnology and Oceanography Methods **7**, 157–168 (2009)
21. Tsikrika, T., García Seco de Herrera, A., Müller, H.: Assessing the scholarly impact of ImageCLEF. In: CLEF 2011. pp. 95–106. Springer Lecture Notes in Computer Science (LNCS) (sep 2011)
22. Tsikrika, T., Larsen, B., Müller, H., Endrullis, S., Rahm, E.: The scholarly impact of CLEF (2000–2009). In: Information Access Evaluation. Multilinguality, Multimodality, and Visualization, pp. 1–12. Springer (2013)
23. Villegas, M., Paredes, R.: Overview of the ImageCLEF 2012 scalable web image annotation task. In: CLEF Working Notes (2012)
24. Villegas, M., Paredes, R.: Overview of the ImageCLEF 2014 scalable concept image annotation task. In: CLEF Working Notes. pp. 308–328. Citeseer (2014)
25. Villegas, M., Paredes, R., Thomee, B.: Overview of the ImageCLEF 2013 scalable concept image annotation subtask. In: CLEF Working Notes (2012)