Organiser Team at ImageCLEFlifelog 2020: A Baseline Approach for Moment Retrieval and Athlete Performance Prediction using Lifelog Data

Tu-Khiem Le[1], Van-Tu Ninh[1], Liting Zhou[1], Minh-Huy Nguyen-Ngoc[5],
Huu-Duc Trinh[5], Nguyen-Hien Tran[5], Luca Piras[2], Michael Riegler[3], Pål
Halvorsen[3], Mathias Lux[4], Minh-Triet Tran[5], Graham Healy[1], Cathal Gurrin[1],
and Duc-Tien Dang-Nguyen[6]

[1] Dublin City University, Dublin, Ireland
[2] Pluribus One & University of Cagliari, Cagliari, Italy
[3] Simula Research Laboratory, Oslo, Norway
[4] Klagenfurt University, Klagenfurt, Austria
[5] University of Science, VNU-HCM, Ho Chi Minh City, Vietnam
[6] University of Bergen, Bergen, Norway

Abstract. For the LMRT task at ImageCLEFlifelog 2020, LIFER 3.0, a new version of the LIFER system with improvements in the user interface and system affordance, is used and evaluated via feedback from a user experiment. In addition, since both tasks share a common dataset, LIFER 3.0 borrows some features from the LifeSeeker system deployed for the Lifelog Search Challenge; which are free-text search, visual similarity search and elastic sequencing filter. For the SPLL task, we proposed a naive solution by capturing the rate of change in running speed and weight, then obtain the target changes for each subtask using average computation and linear regression model. The results presented in this paper can be used as comparative baselines for other participants in the ImageCLEFlifelog 2020 challenge.

1 Introduction

Recent advances in low-cost sensing technologies have resulted in a rapid increase in the volume of digital records (i.e. pictures, videos, audio clips) generated by personal devices such as smartphones, cameras, or wearble devices. This has resulted in a need for efficient management systems to organise and retrieve information from such archives. As a result, many efforts have been made to put together lifelog data and state-of-the-art methods to develop interactive search engines to serve this purpose, which are evaluated via various benchmarking challenges, namely NTCIR [11–13], LSC [14, 15] and ImageCLEFlifelog [4–6]

In the 2020 edition of ImageCLEF2020 [16], the Lifelog Moments Retrieval Task [24] (LMRT) of ImageCLEF2020lifelog challenge has utilised a bigger dataset with 114 days of lifelog data, which is the same as the dataset used in the Lifelog Search Challenge 2020 [14]. The ultimate goal of LMRT is to retrieve a number of relevant moments which match a given query. Besides, another brand-new lifelog task - Sport Lifelog Performance (SPLL) - was proposed with the aim of predicting the expected performance of athletes after they trained for a sporting event. The SPLL data was gathered from 16 different people who trained for a sporting event for approximately six months. The data was collected using three different approaches including wearable devices (Fitbit Tracker, Fitbit Versa) for biometrics data recording (heart rate, calories, speed, pace, running distance, etc.), Google Forms for self-reporting, and PMSYS for subjective wellbeing, injuries and training load. The SPLL task was split into three small subtasks as follows:

1. Predict the change in running speed given by the change in seconds used per km (kilometer speed) from the initial run to the run at the end of the reporting period.
2. Predict the change in weight since the beginning of the reporting period to the end of the reporting period in kilograms.
3. Predict the change in weight from the beginning of February to the end of the reporting period in kilograms using the images.

For the LMRT task, we inherited the design of LifeSeeker [20, 21] with free-text search, external visual concepts detector and temporal exploration using elastic sequencing. We introduce changes to the system's interface in order to tackle the LMRT task and conducted a user study to gain more insights into the performance of the search engine. Moreover, we give an overview of our search engine and how the user study is set up and analyse the results on the LMRT task. For the SPLL task, we provide basic approaches and baseline solutions to predict the expected performance of the athletes, including the change in running speed and weight from the recorded data of Fitbit device and food images only.

2 Related Work

MyLifeBits [10] was a pioneering system which enabled interation between end users and lifelog data using a basic interactive retrieval mechanism. This interation was then enhanced by the work of Doherty et al [9] which allows human to create faceted queries, making it one of the first multimodal interactive lifelog retrieval systems. Due to the increasing attention on lifelogging, many lifelog search engines have been developed, which escalates the need to have a fair comparison among systems. Hence, the availability of the annual challenges such as NTCIR Lifelog Task [11–13], Lifelog Search Challenge [14,15] and ImageCLEF-lifelog [4–6, 24] have successfully facilitated the comparative evaluation of retrieval systems while also supporting and facilitating reserachers to make progress in a shared and collaborative environment.

Considering specifically the LMRT task of ImageCLEFlifelog 2019 [6], nine teams took part in the challenge with a wide variety of approaches to address the problem in automatic manner and interactive manner. There was a general trend that most of the teams extended the provided visual concept annotations by utilising various concept dectectors [19, 25, 26, 28], which was believed to enhance retrieval performance. For automatic runs, the retrieval approach employed by most teams was similar, which was eliminating low quality images and calculating similarity based on relevant scores [8, 26, 30]. In contrast, for interactive systems, we observed that different variences of the Bag-of-Words model were applied to the whole dataset to generating embeddings which served as the backbone for the search engines [7, 19, 25]. Our system relied on the provided metadata and also integrated additional visual concepts to solve this year challenge. The design of LIFER 3.0 is presented in the following section.

3 Overview of LIFER 3.0- Baseline Interactive Search Engine for the Lifelog Moment Retrieval Task (LMRT)
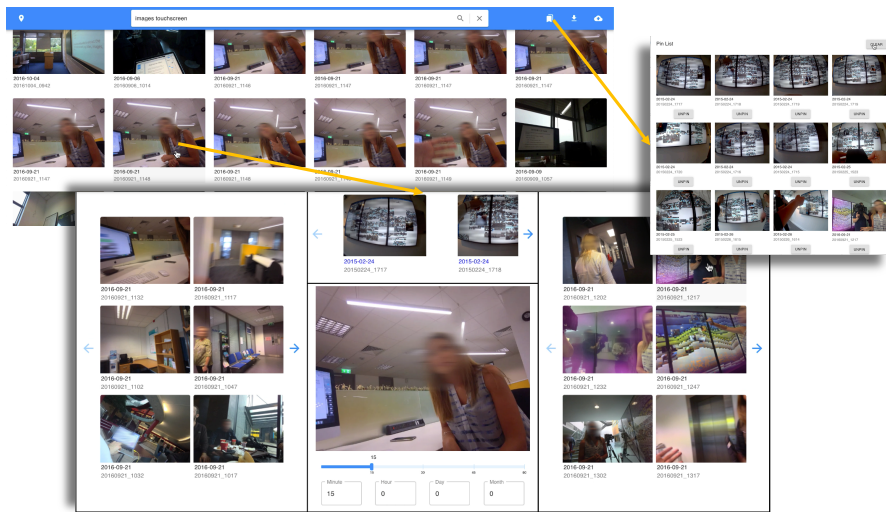

Fig. 1. Changes in the interface of LIFER 3.0. The elastic sequencing filter is simplied and pin list is introduced to the system to manage the selected images.

In this task, we introduce LIFER 3.0 as the baseline interactive retrieval search engine, which is an improved version of the previous baseline systems at ImageCLEFlifelog challenge [25, 29]. LIFER 3.0 inherited the advancements made in LifeSeeker [20, 21] - interactive retrieval system at LSC, which were recently implemented for the NTCIR14 Lifelog Task as its baseline system [23].

Although the LMRT task and LSC challenge share the same dataset containing 114 days with 191,439 lifelogging images and corresponding metadata (biometrics, location and GPS, human activity, visual concepts and annotations), the ultimate goal of each task is different. LSC aims at retriving a single image that perfectly fits the given narrative while LMRT expects the result to be a ranked list of relevant images (moments), which match the description and cover a wide range of moments. Therefore, by ultilising LifeSeeker for this task, we want to evaluate the performance of this search engine in terms of relevant images (precision) and moments coverage (recall).

Our system as described in [21] provides a free-text search mechanism to ensure the simplicity for users (even novice users) to learn and use the search engine. The underlying process includes parsing the input text-query into various lexemes and mapping into different part-of-speech tags (POS tags) using a natural language processor (nltk [22]). This enables us to perform term matching between the query and index with a higher granularity to produce a ranked list of target images.

The index for the system is initialised with the similar approach using nltk where each image is converted into a collection of terms organised into multiple fields such as: time, location, visual concepts, etc. We further employed the bottom up attention method [1] which is pre-trained on Visual Genome [17] so as to better tag the images since Visual Genome comes with larger range of object classes and object attributes. Besides, the image annotation is also extended to include any text appearing within them, which was generated using text recognition from CRAFT [2]

In order to adapt LifeSeeker to work for LMRT, we modified the interface of the search engine to let a user quickly preview an image by hovering over it while pressing the button x and select mutiple images for submission by right-clicking them. The selected images are held locally as pinned items and can be viewed and revised before submitting the results of the queries. The elastic sequecing filters from LifeSeeker which display past and future images are simplified and merged into one single filter to minimise the interaction effort required and reduce the searching time. Figure 1 illustrates the changes introduced in the interface of LIFER 3.0.

4   Sport Performance Lifelog Task: A Baseline Approach

In total, we submitted two runs which are different only in subtask 1 as we nominated two approaches for this subtask.

Subtask 1: *Predict the change in running speed given by the change in seconds used per km (kilometer speed) from the initial run to the run at the end of the reporting period.*

In this subtask, we exploit the exercise data recorded from Fitbit Tracker to gain information about exercise activities, running distance, exercise duration to infer the pace, which is the kilometer speed measure in seconds. Therefore,

we filter the list of exercise activities and keep the fields with the information of distance and exercise duration to compute the pace.

In run 1, at first, we compute the change between pace of consecutive filtered activities. Then, we split them into positive and negative changes to compute average for each type of changes and finally sum them. In run 2, we only consider running and treadmill training as they both involve running activities and follow the procedure in run 1 to obtain the sum of positive and negative average changes for the two activities separately. We then train a linear regression model to predict the actual pace change from the pace of running and treadmill training activities.

Subtask 2: *Predict the change in weight since the beginning of the reporting period to the end of the reporting period in kilograms.*

In this subtask, we employ the self-reporting weight to calculate the change between the start of logging period to its end. The approach is the same as in run 1 of subtask 1. We compute the difference between the weight of consecutive rows in self-reporting files. Then we divide the difference based on its sign (positive or negative) to calculate average for each type and finally sum them.

Subtask 3: *Predict the change in weight from the beginning of February to the end of the reporting period in kilograms using the images.*

To predict the weight changes based on food images, we train a Convolutional Neural Network using Inception V3 architecture [27] on the Food-101 [3] dataset to detect the kind of the food in the images. From the name of the food, we search for the calories information of the food in the nutrition database [7] and use it to estimate the weight gain after the athelete had meal.

5 Experiment and Results

5.1 LMRT Task

We carried out a user study using three participants (two novice users and one expert user) for the search task, each accounts for one run in our submission. The expert user is the author of the system while the novice users are the people who has no prior to knowledge of lifelogging and the search engine. The participants are given a brief introduction to lifelogging, lifelog data and an overview of the functionalities of LIFER 3.0. We allowed the participants to freely explore the search engine using the development queries as long as they wish. The experiment started once they are ready and familiar with the search engine.

In the experiment, all LMRT test queries are presented to the participants with a time limit of five minutes per query. However, there was no time limit for reading a query's description and narrative, so the participants could spend as long as they wish to understand the query before beginning the search process. We provided no clarification or guidance to the user during this user study. Once they finished their search task, there was a follow-up questionaire to be filled to get their opinion about LIFER 3.0. The list of questions are derived from the User Experience Questionnaire - QEU [18]

---

[7] http://nutritionix.com

Table 1. Submitted Runs for LMRT task

RunID	P@10	CR@10	F1@10
Run 1 (Novice)	0.19	0.31	0.21
Run 2 (Novice)	0.23	**0.44**	0.27
Run 3 (Expert)	**0.36**	0.38	**0.32**

Table 1 displays the result of our three runs where each run is generated by one participant. Among the three runs, our system achieved a precision score of 36% on Run 3, while getting 44% in terms of recall on Run 2 and reached the result of 32% in overall F1-Score on Run 3. As can be seen from the table, the expert user tends to perform better than novice users. Nonetheless, the gap between the scores of novice users and expert user is not high, which indicates that LIFER 3.0 might be generalised enough for any user to perform the search task.

Moreover, we further analysed the precision and recall scores across multiple cut-off positions. As illustrated in Figure 2, LIFER 3.0 achieved a high precision at top 5 images and dropped gradually. This means that users are able to select correct images as soon as the results are presented to them. For recall score, we observed a flat curve in the submission of most participants, which implies that the number of distinct moments the users select doesn't change regardless of the cut-off positions. This happens due to the participants not trying to select all relevant images nor look for other similar moments. Therefore, it is possible to boost the overall scores if we further apply a post-processing and re-ranking to append simliar target images to the list submitted by the users.
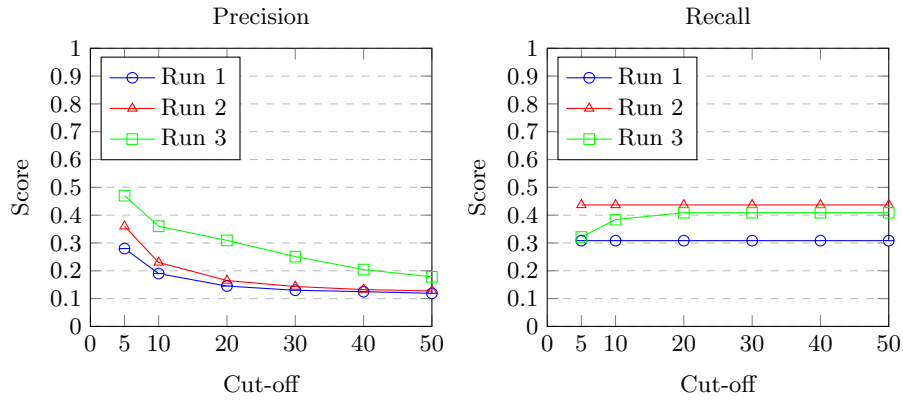
Fig. 2. Precision and Recall of LIFER 3.0 across multiple cut-off positions

From the user questionnaire, we note that the pragmatic quality and hedonic quality demonstrated some proficiences in some criteria while indicating some

areas of improvement that we should work on in the future. In terms of pragmatic quality, LIFER 3.0 is very easy (+2.0) and slightly supportive (+1.0) to the users, but it is also moderately inefficient (-1.0) and confusing (-1.0). The ustilisation of free-text search is probably the main factor which contributes to the ease of using the system for the participants. Moving to the hedonic quality, LIFER 3.0 is quite exciting and interesting to the users and they see it as a bit usual (-0,7) system which is half way between conventional (0.0) and inventive (0.0).

Based on the evaluation results and users' feedback, we identified some concrete actions that need to be realised to improve our system:

- Continue to work on search algorithm to increase the system's efficiency by performing better matching between queries and data and minising the execution time.
- Revise the user interface to present result in a clear and logical manner. Some instruction will be added to serve as system guide in order to lower the confusion.

Table 2. Quality feedback of LIFER 3.0

Item	Mean	Variance	Std. Dev.	Negative	Positive	Scale
1	0.3	0.3	0.6	Obstructive	Supportive	
2	2.0	0.0	0.0	Complicated	Easy	Pragmatic Quality
3	-1.0	1.0	1.0	Inefficient	Efficient	
4	-1.0	1.0	1.0	Confusing	Clear	
5	1.0	0.0	0.0	Boring	Exciting	
6	1.0	0.0	0.0	Not Interesting	Interesting	Hedonic Quality
7	0.0	0.0	0.0	Conventional	Inventive	
8	-0.7	1.3	1.3	Usual	Leading Edge	

5.2 SPLL Task

Table 3. Results of submitted runs in SPLL task

Run	Primary Score	Secondary Score
RUN1	0.47	313.30
RUN2	0.41	203.10

As illustrated in the Table 3, there is a small difference between two submitted runs in terms of primary score. However, we observed a large gap in the secondary score between run 1 and run 2. The difference is the result of two approaches in tackling the subtask 1 - pace change estimation. The average computation approach captures the direction of pace changes better as it takes the rate of

positive and negative change of each into account. Despite of that, it fails to estimate the change in seconds when dealing with multiple types of exercise and training activities, which lower the secondary score. The linear regression model, in contrast, provides a better estimation since it learns to combine the changes in running sessions and treadmill sessions. The detailed score for each subtask of the our baseline approaches are presented in Table 4.

Table 4. Detailed results of each task of our runs in SPLL task.

Run	Task	Primary Score	Secondary Score
RUN1	1	0.6	302.8
	2	0.4	8.5
	3	0.5	2.0
RUN2	1	0.4	192.6
	2	0.4	8.5
	3	0.5	2.0

6 Conclusion

In this paper, we present a baseline solution for both challenges in ImageCLE-Flifelog 2020. For SPLL task, to predict whether the change in running time per kilometer and weight after training is improvement or deterioration, we proposed a basic solution by accumulating the difference between consecutive targeted values, then compute the average of positive and negative difference separately, and finally sum them. For LMRT task, we introduced a baseline interactive search engine which is derived from the LifeSeeker search engine from Lifelog Search Challenge with three main features which are the free-text search, visual similarity exploration and temporal views using elastic sequencing. We have successfully established a user study and drawn many insights from the experiment in terms of interaction and performance through the task's evaluation and users' quality feedback. For the future development, we are aiming to improve the search interface to present the retrieval results efficiently and continuing to work on the core functionalities of search engine to perform better matching between queries and dataset in order to boost the overall search accuracy.

7 Acknowledgement

References

1. Anderson, P., He, X., Buehler, C., Teney, D., Johnson, M., Gould, S., Zhang, L.: Bottom-up and top-down attention for image captioning and visual question answering. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 6077–6086 (2018)
2. Baek, Y., Lee, B., Han, D., Yun, S., Lee, H.: Character region awareness for text detection. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) pp. 9357–9366 (2019)
3. Bossard, L., Guillaumin, M., Van Gool, L.: Food-101 – mining discriminative components with random forests. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) Computer Vision – ECCV 2014. pp. 446–461. Springer International Publishing, Cham (2014)
4. Dang-Nguyen, D.T., Piras, L., Riegler, M., Boato, G., Zhou, L., Gurrin, C.: Overview of imagecleflifelog 2017: Lifelog retrieval and summarization. In: CLEF (2017)
5. 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: CLEF (2018)
6. Dang-Nguyen, D.T., Piras, L., Riegler, M., Zhou, L., Lux, M., Tran, M.T., Le, T.K., Ninh, V.T., Gurrin, C.: Overview of imagecleflifelog 2019: Solve my life puzzle and lifelog moment retrieval. In: CLEF (2019)
7. Dao, M.S., Vo, A.K., Phan, T.D., , Zettsu, K.: BIDAL@imageCLEFlifelog2019: The Role of Content and Context of Daily Activities in Insights from Lifelogs. In: CLEF2019 Working Notes. CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, Lugano, Switzerland (2019)
8. Dogariu, M., Ionescu, B.: Multimedia Lab @ ImageCLEF 2019 Lifelog Moment Retrieval Task. In: CLEF2019 Working Notes. CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, Lugano, Switzerland (2019)
9. Doherty, A.R., Pauly-Takacs, K., Caprani, N., Gurrin, C., Moulin, C.J.A., O'Connor, N.E., Smeaton, A.F.: Experiences of aiding autobiographical memory using the sensecam. Human–Computer Interaction 27(1-2), 151–174 (2012)
10. Gemmell, J., Bell, G., Lueder, R.: Mylifebits: a personal database for everything. Commun. ACM 49(1), 88–95 (2006), http://dblp.uni-trier.de/db/journals/cacm/cacm49.htmlGemmellBL06
11. Gurrin, C., Joho, H., Hopfgartner, F., Zhou, L., Albatal, R.: Overview of ntcir-12 lifelog task. In: NTCIR (2016)
12. Gurrin, C., Joho, H., Hopfgartner, F., Zhou, L., Gupta, R., Albatal, R., Nguyen, D.T.D.: Overview of ntcir-13 lifelog-2 task (2017)
13. Gurrin, C., Joho, H., Hopfgartner, F., Zhou, L., Ninh, V.T., Le, T.K., Albatal, R., Dang-Nguyen, D.T., Healy, G.: Overview of the ntcir-14 lifelog-3 task (2019)
14. Gurrin, C., Le, T.K., Ninh, V.T., Dang-Nguyen, D.T., Jónsson, B.T., Lokoš, J., Hürst, W., Tran, M.T., Schöffmann, K.: Introduction to the third annual lifelog search challenge (lsc'20). In: Proceedings of the 2020 International Conference on Multimedia Retrieval. p. 584–585. ICMR '20, Association for Computing Machinery, New York, NY, USA (2020), https://doi.org/10.1145/3372278.3388043
15. Gurrin, C., Schoe, K., Mann, Joho, H., Leibetseder, A., Zhou, L., Duane, A., Dang-Nguyen, D.T., Riegler, M.A., Piras, L., Tran, M.T., Loko, J., Hürst, W.: Paper comparing approaches to interactive lifelog search at the lifelog search challenge (lsc 2018) (2019)

16. Ionescu, B., Müller, H., Péteri, R., Abacha, A.B., Datla, V., Hasan, S.A., Demner-Fushman, D., Kozlovski, S., Liauchuk, V., Cid, Y.D., Kovalev, V., Pelka, O., Friedrich, C.M., de Herrera, A.G.S., Ninh, V.T., Le, T.K., Zhou, L., Piras, L., Riegler, M., l Halvorsen, P., Tran, M.T., Lux, M., Gurrin, C., Dang-Nguyen, D.T., Chamberlain, J., Clark, A., Campello, A., Fichou, D., Berari, R., Brie, P., Dogariu, M., Ştefan, L.D., Constantin, M.G.: Overview of the ImageCLEF 2020: Multimedia retrieval in lifelogging, medical, nature, and internet applications. In: Experimental IR Meets Multilinguality, Multimodality, and Interaction. Proceedings of the 11th International Conference of the CLEF Association (CLEF 2020), vol. 12260. LNCS Lecture Notes in Computer Science, Springer, Thessaloniki, Greece (September 22-25 2020)

17. Krishna, R., Zhu, Y., Groth, O., Johnson, J., Hata, K., Kravitz, J., Chen, S., Kalantidis, Y., Li, L.J., Shamma, D.A., Bernstein, M., Fei-Fei, L.: Visual genome: Connecting language and vision using crowdsourced dense image annotations (2016), https://arxiv.org/abs/1602.07332

18. Laugwitz, B., Held, T., Schrepp, M.: Construction and evaluation of a user experience questionnaire. In: Holzinger, A. (ed.) HCI and Usability for Education and Work. pp. 63–76. Springer Berlin Heidelberg, Berlin, Heidelberg (2008)

19. Le, N.K., Nguyen, D.H., Nguyen, V.T., Tran, M.T.: Lifelog Moment Retrieval with Advanced Semantic Extraction and Flexible Moment Visualization for Exploration. In: CLEF2019 Working Notes. CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, Lugano, Switzerland (2019)

20. Le, T.K., Ninh, V.T., Dang-Nguyen, D.T., Tran, M.T., Zhou, L., Redondo, P., Smyth, S., Gurrin, C.: Lifeseeker: Interactive lifelog search engine at lsc 2019. In: Proceedings of the ACM Workshop on Lifelog Search Challenge. p. 37–40. LSC '19, Association for Computing Machinery, New York, NY, USA (2019), https://doi.org/10.1145/3326460.3329162

21. Le, T.K., Ninh, V.T., Tran, M.T., Nguyen, T.A., Nguyen, H.D., Zhou, L., Healy, G., Gurrin, C.: Lifeseeker 2.0: Interactive lifelog search engine at lsc 2020. In: Proceedings of the Third Annual Workshop on Lifelog Search Challenge. p. 57–62. LSC '20, Association for Computing Machinery, New York, NY, USA (2020), https://doi.org/10.1145/3379172.3391724

22. Loper, E., Bird, S.: Nltk: The natural language toolkit. In: Proceedings of the ACL-02 Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics - Volume 1. p. 63–70. ETMTNLP '02, Association for Computational Linguistics, USA (2002), https://doi.org/10.3115/1118108.1118117

23. Ninh, V.T., Le, T.K., Zhou, L., Healy, G., Tran, M.T., Dang-Nguyen, D.T., Smyth, S., Gurrin, C.: A Baseline Interactive Retrieval Engine for the NTICR-14 Lifelog-3 Semantic Access Task. In: The Fourteenth NTCIR conference (NTCIR-14) (2019)

24. Ninh, V.T., Le, T.K., Zhou, L., Piras, L., Riegler, M., l Halvorsen, P., Tran, M.T., Lux, M., Gurrin, C., Dang-Nguyen, D.T.: Overview of ImageCLEF Lifelog 2020:Lifelog Moment Retrieval and Sport Performance Lifelog. In: CLEF2020 Working Notes. CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, Thessaloniki, Greece (September 22-25 2020)

25. Ninh, V.T., Le, T.K., Zhou, L., Piras, L., Riegler, M., Lux, M., Tran, M.T., Gurrin, C., Dang-Nguyen, D.T.: Lifer 2.0: Discovering personal lifelog insights using an interactive lifelog retrieval system. In: CLEF (2019)

26. Ribeiro, R., Neves, A.J.R., Oliveira, J.L.: UAPTBioinformatics working notes at ImageCLEF 2019 Lifelog Moment Retrieval (LMRT) task. In: CLEF2019 Work-

ing Notes. CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, Lugano, Switzerland (2019)

27. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the inception architecture for computer vision. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE (jun 2016),

28. Taubert, S., Kahl, S.: Automated Lifelog Moment Retrieval based on Image Segmentation and Similarity Scores. In: CLEF2019 Working Notes. CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, Lugano, Switzerland (2019)

29. Zhou, L., Piras, L., Riegler, M., Lux, M., Dang-Nguyen, D.T., Gurrin, C.: An interactive lifelog retrieval system for activities of daily living understanding. In: CLEF (2018)

30. Zhou, P., Bai, C., Xia, J.: ZJUTCVR Team at ImageCLEFlifelog2019 Lifelog Moment Retrieval Task. In: CLEF2019 Working Notes. CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, Lugano, Switzerland (2019)