Researching Responsible Natural Language Processing

Session 14: Scientific Writing: Abstracts

Frank Keller

1 November 2024

School of Informatics
University of Edinburgh
keller@inf.ed.ac.uk

Overview

Importance

Content

Organization

Reading: Alley (2018), pp. 139–143, Zobel (2014), p. 57.

Please also look at Alley's web site, which has a lot of videos and additional materials: https://www.craftofscientificwriting.com/

Recap

In previous sessions on scientific writing, we talked about:

- audience, purpose, occasion
- precision and clarity
- structure

In this session, we will apply what we've learned to one of the most important parts of a paper: **the abstract.**

Importance

Why is the Abstract Important?

For readers:

- find out what the paper is about (beyond the information in the title)
- get a summary of the most important findings
- decide whether they want to read the paper or not

For reviewers:

- find out whether the paper is within their area of expertise
- decide whether they want to review the paper or not
- form a first opinion about the quality of the paper

Many readers will **only read the title and the abstract.** So this is the one chance to get your message across!

Why is the Abstract Important?

The abstract can also be a tool for the writer:

- helps you decide what your most important points are
- helps you clarify the overall argumentation of the paper
- provides a way of repeating important information
- allows you to influence who will read (and review!) the paper

Content

Content of the Abstract

According to Alley (who calls it summary), the abstract should:

- contain the most important points of the paper
- contain **only** the important points
- only include material that occurs elsewhere in the paper (verbatim or paraphrased)
- be self-contained, i.e., the reader should be able to understand the abstract without having to read anything else
- this means unusual terms, techniques, etc. need to be explained in the abstract
- don't assume all readers will be specialists in the topic of the paper; assume a broad readership.

Content of the Abstract

Alley distinguishes:

- **informative summary:** describes the most important results of a paper;
- **descriptive summary:** states what kind of information the paper provides (like a table of contents), but doesn't give the actual results.

The abstract of a conference or journal paper is a mixture of both: it provides signposting (which information to expect in the paper), but also summarizes the results.

Content of the Abstract

Zobel offers the following practical advice:

- the abstract is typically a single paragraph of about 50–200 words
- it presents a summary of the paper's aims, scope, and conclusions
- do not use acronyms, mathematics, abbreviations, citations (the abstract should be self-contained!)
- be as specific as possible (instead of *we improve the state of the art*, write things like *we improve the state of the art by 3.5%*)
- but only include **important** details.

Over to You

Exercise 1: Evaluating an Abstract

Look at the abstracts on the next pages. Investigate the following questions:

1. Can you identify a structure that these abstracts follow? Is Zobel right?
2. Which audience do the abstracts target? Is Alley right (general reader)?
3. Are the abstracts self-contained, i.e., they are understandable independent of the paper?
4. Do they use acronyms, mathematics, abbreviations, citations (Alley and Zobel)?
5. Is enough detail provided, and is all the detail important?

Both abstracts are from papers that appeared at ACL 2024.

Exercise 1: Evaluating an Abstract

Abstract of Lee et al. (2024):

Most existing image captioning evaluation metrics focus on assigning a single numerical score to a caption by comparing it with reference captions. However, these methods do not provide an explanation for the assigned score. Moreover, reference captions are expensive to acquire. In this paper, we propose FLEUR, an explainable reference-free metric to introduce explainability into image captioning evaluation metrics. By leveraging a large multimodal model, FLEUR can evaluate the caption against the image without the need for reference captions, and provide the explanation for the assigned score. We introduce score smoothing to align as closely as possible with human judgment and to be robust to user-defined grading criteria. FLEUR achieves high correlations with human judgment across various image captioning evaluation benchmarks and reaches state-of-the-art results on Flickr8k-CF, COMPOSITE, and Pascal-50S within the domain of reference-free evaluation metrics. Our source code and results are publicly available at: https://github.com/Yebin46/FLEUR.

Exercise 1: Evaluating an Abstract

Abstract of Chen et al. (2024):

Fine-grained vision-language models (VLM) have been widely used for inter-modality local alignment between the predefined fixed patches and textual words. However, in medical analysis, lesions exhibit varying sizes and positions, and using fixed patches may cause incomplete representations of lesions. Moreover, these methods provide explainability by using heatmaps to show the general image areas potentially associated with texts rather than specific regions, making their explanations not explicit and specific enough. To address these issues, we propose a novel Adaptive patch-word Matching (AdaMatch) model to correlate chest X-ray (CXR) image regions with words in medical reports and apply it to CXR-report generation to provide explainability for the generation process. AdaMatch exploits the fine-grained relation between adaptive patches and words to provide explanations of specific image regions with corresponding words. To capture the abnormal regions of varying sizes and positions, we introduce an Adaptive Patch extraction (AdaPatch) module to acquire adaptive patches for these regions adaptively. Aiming to provide explicit explainability for the CXR-report generation task, we propose an AdaMatch-based bidirectional LLM for Cyclic CXR-report generation (AdaMatch-Cyclic). It employs AdaMatch to obtain the keywords for CXR images and 'keypatches' for medical reports as hints to guide CXR-report generation. Extensive experiments on two publicly available CXR datasets validate the effectiveness of our method and its superior performance over existing methods. Source code will be released.

Organization

Organization of the Abstract

Zobel suggests to start by writing one sentence on each of the following:

1. A general statement introducing the broad research area.
2. An explanation of the specific problem to be solved.
3. A review of existing solutions and their limitations.
4. An outline of the proposed new solution.
5. A summary of how the solution was evaluated and the result of the evaluation.

So you start with five sentences, but then you can add additional sentences, re-write the ones you have, merge them, etc.

Organization of the Abstract

My own experience shows:

- longer documents may require longer abstracts: the abstract of a journal paper is somewhat longer than that of a conference paper
- the abstract of a PhD thesis is typically a whole page; it should summarize each (content) chapter
- abstracts can contain sentences extracted from the main body of the text (you may need to edit them for coherence)
- but: it is sometimes a good strategy to write the abstract **before** writing the paper – helps planning the overall argumentation, deciding what to focus on
- and then once the paper is finished, you need to completely re-write the abstract!

Over to You

Fine-Grained Image-Text Alignment in Medical Imaging Enables Explainable Cyclic Image-Report Generation

Wenting Chen[1] **Linlin Shen**[3] **Jingyang Lin**[4] **Jiebo Luo**[4]
Xiang Li[5*] **Yixuan Yuan**[2*]
[1]City University of Hong Kong [2]The Chinese University of Hong Kong
[3]Shenzhen University [4]University of Rochester
[5]Massachusetts General Hospital and Harvard Medical School
[1]wentichen7-c@my.cityu.edu.hk [2]yxyuan@ee.cuhk.edu.hk [3]llshen@szu.edu.cn
[4]{jluo@cs, jlin81@ur}.rochester.edu [5]xli60@mgh.harvard.edu

Abstract

Fine-grained vision-language models (VLM) have been widely used for inter-modality local alignment between the predefined fixed patches and textual words. However, in medical analysis, lesions exhibit varying sizes and positions, and using fixed patches may cause incomplete representations of lesions. Moreover, these methods provide explainability by using heatmaps to show the general image areas potentially associated with texts rather than specific regions, making their explanations not explicit and specific enough. To address these issues, we propose a novel Adaptive patch-word Matching (AdaMatch) model to correlate chest X-ray (CXR) image regions with words in medical reports and apply it to CXR-report generation to provide explainability for the generation process. AdaMatch exploits the fine-grained relation between adaptive patches and words to provide explanations of specific image regions with corresponding words. To capture the abnormal regions of varying sizes and positions, we introduce an Adaptive Patch extraction (AdaPatch) module to acquire adaptive patches for these regions adaptively. Aiming to provide explicit explainability for the CXR-report generation task, we propose an AdaMatch-based bidirectional LLM for Cyclic CXR-report generation (AdaMatch-Cyclic). It employs AdaMatch to obtain the keywords for CXR images and 'keypatches' for medical reports as hints to guide CXR-report generation. Extensive experiments on two publicly available CXR datasets validate the effectiveness of our method and its superior performance over existing methods.

1 Introduction

Inter-modality alignment, such as vision and language, has been an important task with growing interests in the field of computer vision, especially with the recent advancement in representation learning (Radford et al., 2021). Technologies like contrastive learning and self-supervised learning have dramatically improved state-of-the-art alignment performance. Recent vision-language models (VLMs) demonstrate two approaches: global contrastive alignment, which integrates images and texts at a global level (Radford et al., 2021; Jia et al., 2021; Jang et al., 2023; Wang et al., 2023; Yang et al., 2022), and local alignment, focusing on detailed connections between visual objects and textual words (Chen et al., 2020a; Li et al., 2020b,a; Zhan et al., 2021; Kim et al., 2021; Yao et al., 2021), as illustrated in Fig. 1.

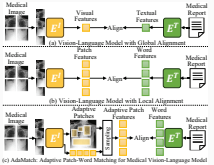

Figure 1: Current vision-language models (VLM) achieve (a) global alignment and (b) local alignment by matching overall visual with textual features, and aligning patches with word features, respectively. (c) To exploit the relation between textual words and abnormal patches with varied sizes, our AdaMatch obtains adaptive patch features and aligns them with word features.

Current VLMs with local alignment either adopt the pre-trained object detector to extract region-of-interest (ROI) features from images and match the corresponding object features with textual words (Chen et al., 2020a; Li et al., 2020b,a; Zhan et al., 2021), or align the visual token from each

*Xiang Li and Yixuan Yuan are corresponding authors.

Exercise 2: Shortening an Abstract

In addition to the problems already discussed, the abstract of Chen et al. (2024) is simply too long!

- It takes up a large chunk of the first page (254 words).
- Not all readers will have the patience for this.
- And it's space you could use elsewhere in the paper! (Remember, abstracts shouldn't contain new information.)

Your task: shorten the abstract to 125 words.

Exercise 2: Shortening an Abstract

Let's ask GPT: "Could you please shorten the following abstract to 125 words: [...]"

Fine-grained vision-language models (VLMs) often align fixed image patches with textual words, but this approach may miss critical details in medical images, where lesions vary in size and position. To improve this, we propose AdaMatch, an Adaptive patch-word Matching model, for correlating regions in chest X-rays (CXRs) with words in medical reports, enhancing explainability. AdaMatch uses adaptive patches, captured by our Adaptive Patch extraction (AdaPatch) module, to represent abnormal regions accurately. For CXR-report generation, we introduce AdaMatch-Cyclic, a bidirectional model that utilizes AdaMatch to identify keywords and "keypatches" as guides. Extensive testing on two public CXR datasets demonstrates AdaMatch's superior performance and effectiveness in generating detailed, interpretable reports. Source code will be made available.

Exercise 3: Writing Your Own Abstract

Write a draft of the abstract of your report for the method study. Use Zobel's guidelines and write one sentence each for:

1. A general statement introducing the broad research area.
2. An explanation of the specific problem to be solved.
3. A review of existing solutions and their limitations.
4. An outline of the proposed new solution.
5. A summary of how the solution was evaluated and the result of the evaluation.

Try to do this as a homework (unassessed).

References

Alley, Michael. 2018. *The Craft of Scientific Writing*. Springer, New York, NY, 4 edition.

Chen, Wenting, Linlin Shen, Jingyang Lin, Jiebo Luo, Xiang Li, and Yixuan Yuan. 2024. Fine-grained image-text alignment in medical imaging enables explainable cyclic image-report generation. In Lun-Wei Ku, Andre Martins, and Vivek Srikumar, editors, *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, Bangkok, Thailand, pages 9494–9509.

Lee, Yebin, Imseong Park, and Myungjoo Kang. 2024. FLEUR: An explainable reference-free evaluation metric for image captioning using a large multimodal model. In Lun-Wei Ku, Andre Martins, and Vivek Srikumar, editors, *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, Bangkok, Thailand, pages 3732–3746.

Zobel, Justin. 2014. *Writing for Computer Science*. Springer, New York, NY, 3 edition.