NewsImages in MediaEval 2026: Automated Image Recommendations with Retrieval and Generation Techniques for News Articles Thumbnails

Lucien Heitz[1,*], Bruno N. Sotic[2], Ali A. Katamjani[3], Qi Bi[3], Bram Bakker[4], Luca Rossetto[5] and Jaap Kamps[2]

[1]*Informfully, Zurich, Switzerland*
[2]*University of Amsterdam, Amsterdam, the Netherlands*
[3]*Utrecht University, Utrecht, the Netherlands*
[4]*Radboud University Nijmegen, Amsterdam, the Netherlands*
[5]*Dublin City University, Dublin, Ireland*

Abstract

The NewsImages challenge in MediaEval 2026 focuses on matching news articles with appropriate thumbnails. Participants receive a large set of English-language articles from international publishers. Given the text of a news article, the goal of the challenge is to recommend a fitting image. The overall quality of the image recommendation is evaluated in a crowd-sourced online event where image fit and relevance are assessed. This task overview paper provides details on the challenge setup, the key lessons learned from past iterations, the online evaluation, and the technical aspects of the group submissions.

Keywords

image retrieval, image generation, online news, news recommendation

1. Introduction

Thumbnails for news articles are an important factor in capturing the reader's attention on online news platforms. New approaches to automated image retrieval and generation were shown to compete with, and in some instances even outperform, editorial selection of visuals [1]. This year's challenge task is a seamless continuation of the 2025 edition [2] that systematically explored and compared image retrieval with generation and editorial selection for news thumbnails. NewsImages 2026 places great emphasis on the user perspective. Therefore, we are not only interested in technical papers and contributions to retrieving/generating image recommendations, but also in more qualitative work on user perception or trust.[1] The target audience for the task includes media professionals (editors, newsrooms, and journalists), researchers in information retrieval and generative AI, and students in computer science and media/journalism.

MediaEval'26: Multimedia Evaluation Workshop, June XX–XX, 2026, Amsterdam, the Netherlands and Online
lucien.heitz@informfully.ch (L. Heitz); nadalic.sotic@uva.nl (B. N. Sotic); a.ahmadikatamjani@uu.nl (A. A. Katamjani); q.bi@uu.nl (Q. Bi); bram.bakker2@ru.nl (B. Bakker); luca.rossetto@dcu.ie (L. Rossetto); kamps@uva.nl (J. Kamps)
0000-0001-7987-8446 (L. Heitz); 0009-0003-2122-2235 (B. N. Sotic); 0009-0008-9545-2814 (A. A. Katamjani); 0000-0002-1047-4790 (Q. Bi); 0009-0009-8615-1056 (B. Bakker); 0000-0002-5389-9465 (L. Rossetto); 0000-0002-6614-0087 (J. Kamps)

[1]https://multimediaeval.github.io/editions/2026/tasks/newsimages

2. Task Description

NewsImages 2026 explores matching news articles with automated image recommendations. To that end, participants receive a large set of English-language articles from international news organizations and media outlets. Given the text of a news article, the goal of this task is to build a recommendation pipeline that uses retrieval techniques, generative AI, or a combination of both to output a suitable image for an article thumbnail. We briefly outline the available resources, the dataset, and the final submission. The complementary online README file provides more information and technical details.[2]

Resource Selection Staying true to MediaEval's principles of promoting reproducible research, we ask participants to share their entire image retrieval pipelines and generation workflows (including all parameters and links to models to reproduce the image recommendations). The image retrieval/generation pipeline must not rely on any closed-source APIs or resources that are not publicly available. For image retrieval, we recommend public datasets like *Yahoo-Flickr Creative Commons 100 Million* (YFCC100M)[3] or *Public Domain 12 Million* (PD12M).[4] For image generation, we require the models to be open source and/or open weights.

Challenge Dataset In terms of the training dataset, participants are given a collection of $8,500$ news articles with images (the article text is in English, collected during 2022–2023 by GDELT).[5] Each new item includes the following properties: $article\_id$ (internal ID of news article used for submission), $article\_url$ (URL of the original news article), $article\_title$ (title of the news article, may include lead), $image\_id$ (internal ID of original image for which a copy is shared), and $image\_url$ (URL of the original news image). The article text itself is not shared, but participating groups are free to retrieve it from the original source. The dataset is the same as in NewsImages 2025. The main idea behind reusing it for training purposes is that it allows participating teams to build upon existing workflows and compare their recommendations with past results.[6] There will be a separate test dataset for the evaluation event this year. The format will be the same as the training dataset.

Run Submissions Participating teams are encouraged to explore different image styles (e.g., photo-realistic or caricature), image subjects (e.g., people, places, things), and combinations thereof. Teams can submit multiple runs. However, all the runs must be substantially different from one another, and teams must ensure that each run is properly documented.[7] There are three key factors to consider. First, all image recommendations must be unique and distinct from one another; i.e., no two articles within a run can share the same image recommendation. Second, no two runs can share the same image recommendation for a given news headline. Third, your submitted runs must not include any of the original images shared in the dataset. We will remove any image recommendation from your runs that does not meet these three requirements for the final evaluation. Next, we would like to share key insights and lessons learned from NewsImages 2025 to help shape the workflows.

[2] README available on GitHub: https://github.com/Informfully/Challenges
[3] YFCC100M website: https://www.multimediacommons.org
[4] PD12M on Huggingface:https://huggingface.co/datasets/Spawning/PD12M
[5] GDELT website: https://www.gdeltproject.org
[6] All images and workflows of NewsImages 2025 are available online: https://github.com/Informfully/Challenges/tree/main/newsimages25
[7] For instructions on how to write the accompanying Working Notes Paper, please go to the official MediaEval 2025 website: https://multimediaeval.github.io/editions/2026

3. Experiment Guidelines

The NewsImage iteration of 2025 was the first to systematically compare image retrieval and generation pipelines. Based on these results, we have several key insights for future submissions that we would like to share with participants. Please read the recommendations carefully and try to incorporate them into your workflows.

Hybrid Pipelines Focus on hybrid pipelines that combine the strengths of image retrieval and image generation (e.g., retrieving images when a specific person is mentioned, and generating images when you need a generic picture of a place). This requires building article-specific workflows that you can switch between. Ensure that you clearly explain your architecture and indicate where retrieval and generation occur.

State-of-the-Art Models For image generation, please consider a top-ranking open-weight text-to-image model[8] (e.g., Z-Image) or text-to-video model (e.g., LTX-2).[9] Remember that prompt engineering can make a huge difference with these newer models. Read the official documentation and follow the guidelines. Furthermore, instead of relying solely on the original headline, use a text-to-text model to draft richer prompts, which you can then plug into the text-to-image model. You can think of, e.g., creating specific prompting templates for different article categories or keywords within the title.

Human Evaluation Benchmarking using offline metrics, such as text-image similarity or accuracy, was shown to be only a poor indicator of the final performance in the online evaluation. We recommend involving friends and family to help assess the quality of your image recommendations and to get valuable feedback. Alternatively, you can use a candidate set of image recommendations that you then present to automated judges, e.g., based on a CLIP model, to assess the article-image fit.

Style Preferences Art style and image subjects have a large impact on performance. Try to explore different options. Cartoons and caricatures were generally well-received. They provide a reliable way to generate images that also comply with existing media policies against creating realistic images, thereby helping avoid publishing potentially fake news.[10]

4. Evaluation

The image recommendations of the submitted runs will be evaluated in a crowd-sourced online event. All participating teams, task organizers, and an external jury will take part in the event. Image fit is the primary criterion for evaluating and assessing images. Fit focuses on whether the images accurately capture the article's key attributes, without depicting any elements not present in the article. During the event, participants are presented with a news headline and a list of image recommendations. The fit of each image is rated on a 5-point Likert scale.

[8] Online text-to-image leaderboard: https://artificialanalysis.ai/image/leaderboard/text-to-image?open-weights=true
[9] Online text-to-video leaderboard: https://artificialanalysis.ai/text-to-video/arena?tab=leaderboard-text
[10] A complete collection of all submitted pictures can be found online on GitHub: https://github.com/Informfully/Challenges/tree/main/newsimages25/images

The winning team is determined by the highest average image-fit rating achieved by their best submitted run. We highly recommend participants to adhere to editorial standards and guidelines for image generation.[11] We are particularly interested in non-photorealistic images that do not suggest they accurately represent real events, to avoid misleading or deceiving readers. There will be three baselines: 1) the original images selected by the editors, 2) top-performing runs from NewsImages 2025 [2], and 3) a state-of-the-art text-to-image and text-to-video models.[12] The test dataset for the evaluation event contains around 850 articles from which we draw a random subsample that is then featured in the online survey.[13]

5. Conclusion

NewsImages 2026 builds on the 2025 iteration, which first combined image retrieval and generation techniques to produce thumbnail recommendations based on news headlines. With this continuation, our goal is to leverage techniques and insights from previous submissions to create new automated workflows. Finally, we want to reiterate that the NewsImages challenge focuses on both the technological aspects and the user's perspective. We also look for contributions that assess the impact of recommendations on audiences and/or critically reflect on the requirements, risks, and opportunities presented by these new techniques. In doing so, we encourage a broad range of research fields and disciplines to participate in the challenge and share their results and insights for generating new image recommendations.

Acknowledgments

We thank Martha Larson, Steven Hicks, and Mihai Gabriel Constantin for their feedback and support through organizing this challenge.

Declaration on Generative AI

During the preparation of this task overview paper, the authors used Grammarly for grammar and spelling checks. After using this service, the authors reviewed and edited the text as needed. They take full responsibility for the publication's content.

References

[1] L. Heitz, A. Bernstein, L. Rossetto, An empirical exploration of perceived similarity between news article texts and images, in: Working Notes Proceedings of the MediaEval 2023 Workshop, 2024. URL: http://ceur-ws.org/Vol-3658/paper8.pdf.

[2] L. Heitz, L. Rossetto, B. Kille, A. Lommatzsch, M. Elahi, D.-T. Dang-Nguyen, Newsimages in mediaeval 2025 – comparing image retrieval and generation for news articles, in: Working Notes Proceedings of the MediaEval 2025 Workshop, 2025. URL: https://2025.multimediaeval.com/paper1.pdf.

[11] For an overview of existing policies, please see: https://www.ebu.ch/groups/ai-ethics

[12] The selection of the models will be communicated together with the details of the test set.

[13] We may include a small number of articles from the training dataset to make a comparison with the submissions from previous iterations.