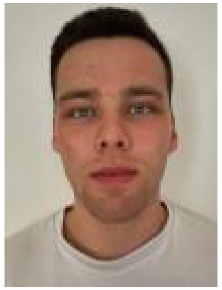
Image Retrieval/Generation for Arguments [Joint Task with ImageCLEF]

Touché'25 Task 3

Maximilian
Heinrich

Johannes
Kiesel



Moritz
Wolter

Martin
Potthast

Benno
Stein

Image Retrieval/Generation for Arguments [Joint Task with ImageCLEF]

Task Description

Scenario: Enhance the impact of arguments.

Task: Given an argument, identify images that effectively convey the argument's premise.

- Participants may either retrieve images from a dataset or generate them using a text-to-image model.

Data:
- 128 arguments across 27 topics
- ca. 32,000 crawled images with corresponding website information and additional metadata, including automatically generated captions

Image Retrieval/Generation for Arguments [Joint Task with ImageCLEF]

Task Description

Scenario: Enhance the impact of arguments.

Task: Given an argument, identify images that effectively convey the argument's premise.

- Participants may either retrieve images from a dataset or generate them using a text-to-image model.

Data:

- 128 arguments across 27 topics
- ca. 32,000 crawled images with corresponding website information and additional metadata, including automatically generated captions

Example

Topic: Public Transportation vs. Private Cars

Claim: Cars make it easy to transport goods and belongings

Image Retrieval/Generation for Arguments [Joint Task with ImageCLEF]

Evaluation

- For each argument, two aspects were identified, and each aspect was rated using the following scale:

  0: Aspect does not convey the claim
  1: Aspect partially conveys the claim
  2: Aspect fully conveys the claim

- For each annotator, the aspect scores were aggregated to derive a single rating for an argument-image pair.

- Final score for an argument-image pair is computed by combining the individual ratings from two annotators.

Image Retrieval/Generation for Arguments [Joint Task with ImageCLEF]

Example Submission

Argument	Retrieval	Generation
Topic: Public Transportation vs. Private Cars **Claim:** Cars make it easy to transport things **Aspects:** car, transport things	 Source: Web	 Source: Stable Diffusion 3.5

Image Retrieval/Generation for Arguments [Joint Task with ImageCLEF]

Example Submission

Argument	Retrieval	Generation
Topic: Public Transportation vs. Private Cars **Claim:** Cars make it easy to transport things **Aspects:** car, transport things		
	Source: Web	Source: Stable Diffusion 3.5

Here both images get a score of two. The two required aspects do not need to be combined in a precise way.

Image Retrieval/Generation for Arguments [Joint Task with ImageCLEF]

Results - Retrieval

Rank	Team	Approach	NDCG@5
1	Baseline	CLIP Image	0.855
2	Infotec+CentroGEO	OpenCLIP Image	0.836
3	Baseline	SBERT Website-Text	0.811
4	Infotec+CentroGEO	MCIP Image	0.794
5	Infotec+CentroGEO	SBERT Image-Text+Caption	0.755
6	CEDNAV–UTB	CLIP Image-Caption	0.236

The 'Approach' column specifies how the embeddings for the images were generated and compared with the arguments. For example, 'CLIP Image' indicates that multimodal CLIP embeddings are employed.

Image Retrieval/Generation for Arguments [Joint Task with ImageCLEF]

Results - Generation

Rank	Team	Approach	NDCG@5
1	Hanuman	Generative Prompt	0.963
2	Baseline	Stable Diffusion 1.0	0.844
3	Baseline	Stable Diffusion 3.5	0.839

Approaches:

- **Generative-Prompt:** Use an LLM to identify key aspects of the argument and compose a tailored image-generation prompt. For generation Stable Diffusion 1.0 is used.
- **Baseline:** Directly use the arguments themselves as the image-generation prompt.

Image Retrieval/Generation for Arguments [Joint Task with ImageCLEF]

Results - Generation

Rank	Team	Approach	NDCG@5
1	Hanuman	Generative Prompt	0.963
2	Baseline	Stable Diffusion 1.0	0.844
3	Baseline	Stable Diffusion 3.5	0.839

Approaches:

- **Generative-Prompt:** Use an LLM to identify key aspects of the argument and compose a tailored image-generation prompt. For generation Stable Diffusion 1.0 is used.
- **Baseline:** Directly use the arguments themselves as the image-generation prompt.

Image generation for arguments produces good results, especially when using carefully crafted custom prompts.

Image Retrieval/Generation for Arguments [Joint Task with ImageCLEF]

Lessons Learned

- Finding suitable images for arguments is challenging; generation often works better for specific arguments than retrieval.
- Retrieval approaches are constrained by the limited scope of available web sources, which tend to emphasize more general arguments.
- The main challenge for generation approaches lies in combining multiple aspects effectively and depicting elements that should not be displayed.