

# MemoriEase Lifelog Retrieval System at NTCIR-18 Lifelog 6 Task

**Engaging Content** Engaging People

#### **Quang-Linh Tran**

ADAPT Centre, Dublin City University, Dublin, Ireland

#### Binh Nguyen

University of Science, Vietnam National University, Ho Chi Minh City, Vietnam Gareth J. F. Jones ADAPT Centre, Dublin City University, Dublin, Ireland Cathal Gurrin ADAPT Centre, Dublin City University, Dublin, Ireland

Taighde Éireann

**Research** Ireland

## Introduction

- Lifelogging: passively collecting and storing personal daily life data.
- MemoriEase system: An interactive system at the Lifelog

## **Results**

• LSAT: MemoriEase finds 83/249 relevance images in KIS and 285 /1746 relevance images in Adhoc. BLIP combined with CQ and text search brings the best performance.

Search Challenge.

- Interactive with a conversational search user interface but adapt to automatic search for Lifelog task at NTCIR.
- Three new points in this version:
  - A comparison of BLIP2 and CLIP embedding models.
  - Use LLM to rewrite query and compare with original query (Concise query).
  - Pseudo relevance feedback for image search vs free-text search for Adhoc query.

## MemoriEase system



• LQAT: 8 over 24 questions are solved.

KIS-LSAT	MAP	MRR	NDCG	R@5	R@20	R@100	P@5	P@20	R@100
CLIP-Q	0.2000	0.4274	0.3279	0.1910	0.4357	0.4940	0.2000	0.1346	0.0638
BLIP-Q	0.2702	0.5242	0.4287	0.2822	0.5035	0.6349	0.2462	0.1462	0.0554
BLIP-CQ	0.2709	0.3425	0.3785	0.2698	0.2847	0.6348	0.2000	0.1192	0.0554

	Adhoc-LSAT	MAP	MRR	NDCG	R@5	R@20	R@100	P@5	P@20	R@100
	CLIP-Q-NoRF	0.0673	0.4391	0.1823	0.0150	0.1223	0.2426	0.2615	0.2462	0.2100
	BLIP-Q-RF	0.0754	0.2827	0.1398	0.0136	0.0382	0.1320	0.2462	0.2358	0.2200
	BLIP-Q-NoRF	0.1661	0.4616	0.3058	0.0550	0.1518	0.3802	0.3692	0.2923	0.1815
	BLIP-CQ-NoRF	0.1805	0.6806	0.3528	0.0598	0.1670	0.4406	0.4308	0.3423	0.2192

Legend: CLIP/BLIP: Embedding model. Q/CQ: Original query vs LLM-rewritten concise query.

NoRF-RF: Text search vs Image search with pseudo relevance feedback.

Table 1: Result on KIS and Adhoc queries in LSAT sub-task.



Figure 1: MemoriEase overall architecture

- A lifelog dataset of 725K images. We remove 120K images of blurry or low-quality, and group images by events before generating descriptions by BLIP2 model.
- Query is rewritten by ChatGPT (consice query).
- Text2Image search: Cosine similarity between image and query embedding.
- Image2Image search (pseudo relevance feedback): Use query to retrieval top 3 images and embed them to vectors to search.

Figure 3: MemoriEase system solved a QA topic in LQAT sub-task

## Conclusion

- MemoriEase shows good performance in the LSAT sub-task and average performance in the LQAT sub-task.
- BLIP is better than CLIP model in encoding lifelog images
- LLM-rewritten queries are useful for Adhoc queries but not
- A RAG Approach for QA task: Classify the question type  $\rightarrow$

Retrieve relevant lifelog data  $\rightarrow$  Prompt for GPT4o with

#### question and contexts $\rightarrow$ Response an answer.

### KIS queries.

- Pseudo relevance feedback with image search for Ad-hoc is not good. It needs human involvement.
- QA task is challenging for LLM.

This research was conducted with the financial support of Science Foundation Ireland at ADAPT, the SFI Research Centre for AI-Driven Digital Content Technology at DCU [13/RC/2106\_P2]. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

