HCMUS at the NTCIR-14 Lifelog-3 Task

Nguyen-Khang Le¹, Dieu-Hien Nguyen¹, Trung-Hieu Hoang¹, Thanh-An Nguyen¹, Thanh-Dat Truong¹, Duy-Tung Dinh¹, Quoc-An Luong¹, Viet-Khoa Vo-Ho¹
Vinh-Tiep Nguyen², Minh-Triet Tran²

¹ University of Science, VNU-HCM, Ho Chi Minh City, Vietnam
² University of Information Technology, VNU-HCM, Vietnam.

INTRODUCTION

Lifelog-3 was the third instance of the lifelog task at NTCIR. At NTCIR-14, the Lifelog-3 task explored three different lifelog data access related challenges. One of the three challenges is Lifelog Semantic Access sub-Task (LSAT) which aims to explore search and retrieval from Lifelogs. In this subtask, the participants had to retrieve a number of specific moments in a lifelogger’s life in response to a query topic. Our proposed system solves this subtask by employing two main features: Data Processing and User Interaction. For Data Processing, we employ models in object detection and scene classification to annotate the lifelog dataset with meaningful metadata. For User Interaction, we aim to design and provide a friendly user interface that enables novice users to interact with the queries and select the result data. In the NTCIR-14 official results, our system has the highest performance.

PROPOSED SYSTEM

Overview

Scene classification

COCO Object Detector

Habit-based Object detector

Scene classification

Habit-based detector

Lifelog retrieval system V2

Web Application

Search panel

RESULTS

The official results of the LSAT

<table>
<thead>
<tr>
<th>Group ID</th>
<th>Run ID</th>
<th>Approach</th>
<th>MAP</th>
<th>P@10</th>
<th>RelRet</th>
</tr>
</thead>
<tbody>
<tr>
<td>NTU</td>
<td>NTU-Run1</td>
<td>Interactive</td>
<td>0.0632</td>
<td>0.2375</td>
<td>293</td>
</tr>
<tr>
<td>NTU</td>
<td>NTU-Run2</td>
<td>Interactive</td>
<td>0.1108</td>
<td>0.3750</td>
<td>464</td>
</tr>
<tr>
<td>NTU</td>
<td>NTU-Run3</td>
<td>Interactive</td>
<td>0.1657</td>
<td>0.6833</td>
<td>407</td>
</tr>
<tr>
<td>DCU</td>
<td>DCU-Run1</td>
<td>Interactive</td>
<td>0.0724</td>
<td>0.1917</td>
<td>556</td>
</tr>
<tr>
<td>DCU</td>
<td>DCU-Run2</td>
<td>Interactive</td>
<td>0.1274</td>
<td>0.2292</td>
<td>1094</td>
</tr>
<tr>
<td>HCMUS</td>
<td>HCMUS-Run1</td>
<td>Interactive</td>
<td>0.3993</td>
<td>0.7017</td>
<td>1444</td>
</tr>
<tr>
<td>QUIK</td>
<td>QUIK-Run1</td>
<td>Automatic</td>
<td>0.0454</td>
<td>0.1958</td>
<td>232</td>
</tr>
<tr>
<td>QUIK</td>
<td>QUIK-Run2</td>
<td>Automatic</td>
<td>0.0454</td>
<td>0.1875</td>
<td>232</td>
</tr>
</tbody>
</table>

Our proposed system has the highest performance in NTCIR-14 result

USER INTERACTION

Sequence view of moments

Filter, reorder results

Search panel

CONCLUSION AND FUTURE WORKS

• There are still some aspects that our system needs to improve. The user still needs to picture the moments to decide what scene category the images should be, and what concepts should be in the images.

• In the future works, we will look into the aspect of natural language semantics to give our system the ability to understand the topic search and suggest more relevant inputs for the user.