# 

SKM4301

**MULTIMEDIA INFORMATION RETRIEVAL**

SEMESTER II 2022-2023

**PART B and PART C: REPORT**

GROUP 5 (IMAGE)

Prepared by:

| **No.** | **Name** | **Matric Number** |
| --- | --- | --- |
| 1 | AMIR NURHAKIM BIN MOHD ZAID | 207092 |
| 2 | ANAS ZULKIFLI BIN MOHD JEFFRY | 206520 |
| 3 | NIK MUHAMMAD ASYRAF BIN NIK ISMAIL | 206630 |
| 4 | MUHAMMAD AMIRUL AFIQ BIN JUMA’AT | 205911 |

# 

# **Task Title**

Content-based Image Retrieval of Specific Vegetables

**Problem Statement**

Build an application of mini-image information retrieval for vegetables which can take a vegetable image as input and retrieve other similar vegetable images from a database.

**Methodology**

Pre-processing:

Extract color space and quantize the values of an input image to store them in an array. For the color space conversion method, the input image extracts R, G, B and negates it to convert the input image into HSB (Hue, Saturation, Brightness) values of 4-bit representation and combines them into a single integer. Then, for an 8-bit integer value, first it converts to a float array and later an integer array of representation of HSB values by using masking and shifting. The objective of this method is to reduce the range of possible values to avoid the complexity of histogram representation. By doing this quantization, a situation where a various image which has a lower-dimensional feature representation or comparing images can be operated.

Indexing:

As for indexing, create a hashmap by iterating over the files in the image database folder. Then, use an MD5 (MD5 is a Message Digest Algorithm 5, a cryptographic hash function) for each file. After that, the code generates the normalized histogram (min-max normalization) of the image for similarity measure later and adds corresponding entry to the hashmap. By using multi-threading, parallelizing the creation process of histogram with MD5 hash is possible therefore it can divide the tasks and wait for all threads to finish for lastly serializing the hashmap. If there is a new image added, then it will go through the same process. As for retrieving an image from the database as output, the distance of the query input image’s histogram will be extracted and compared to all the histogram images in the hashmap. Then, it retrieves the output image which has the nearest distance of histogram.

Similarity Measure:

For similarity measure, Minkowski’s Distance is utilized. Its purpose is to measure the dissimilarity between two histogram distances. The smaller the distance between histograms, the higher similarity of the histograms (images). Minkowski’s Distance works by iterating over the corresponding elements of histogram, calculating squared differences, summing it, taking the square root and lastly returning the result.

**Experimental Setting**

I. Dataset Details

Following the discussion, our group has chosen to focus on the "type of vegetables" for our project. The vegetable dataset obtained from Kaggle, available at the provided link (<https://www.kaggle.com/datasets/misrakahmed/vegetable-image-dataset>), is specifically focused on the classification of different types of vegetables. For the project at hand, the group has chosen to narrow our attention to five specific types of vegetables: carrot, tomato, potato, radish and capsicum.

Within each of these vegetable subsets, the group has selected a total of 10 distinct items, resulting in a dataset containing a total of 50 vegetable images. The dataset includes images of these vegetables captured from various angles and under different conditions, providing a diverse set of visual representations for each vegetable type.

| Carrot |
| --- |
|  |
| Tomato |
|  |
| Potato |
|  |
| Radish |
|  |
| Capsicum |
|  |

II. Evaluation Measurement Used

For our project report, we have chosen to utilize the Precision-Recall Curve as our evaluation measure. The Precision-Recall Curve is a widely employed evaluation metric in information retrieval tasks. It provides valuable insights into the trade-off between precision and recall, allowing us to analyze the performance of our vegetable classification model at various decision boundaries or classification thresholds.

In our vegetable classification project, the Precision-Recall Curve will be particularly valuable for assessing the model's performance, especially considering the potential class imbalance in our dataset and our focus on correctly identifying the vegetable types. By analyzing the curve, we will be able to select an appropriate decision boundary based on our specific requirements. Additionally, we can compute the area under the Precision-Recall Curve to provide a summarized evaluation of the overall performance of our classifier, with higher values indicating better performance.

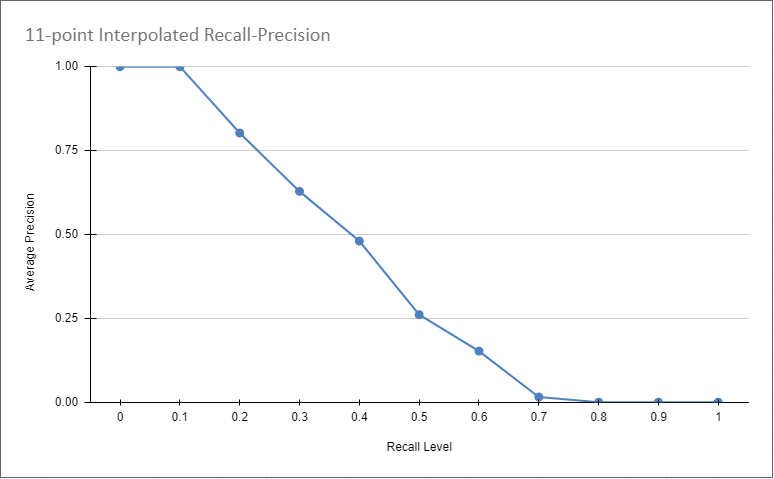
In the Excel sheet attached below, we displayed the result of the Precision-Recall Curve calculated based on our CBIR system.

**Results and Discussion**

The Google Sheet link provided below showcases the detailed table of precision and recall for our system. It presents a comprehensive analysis of how our system performed in terms of precision and recall measures. The calculation considers recall, precision at each recall, the average interpolated recall as well as two 11-point Interpolated Recall Precision graphs generated for the whole system and each category of vegetables dataset.

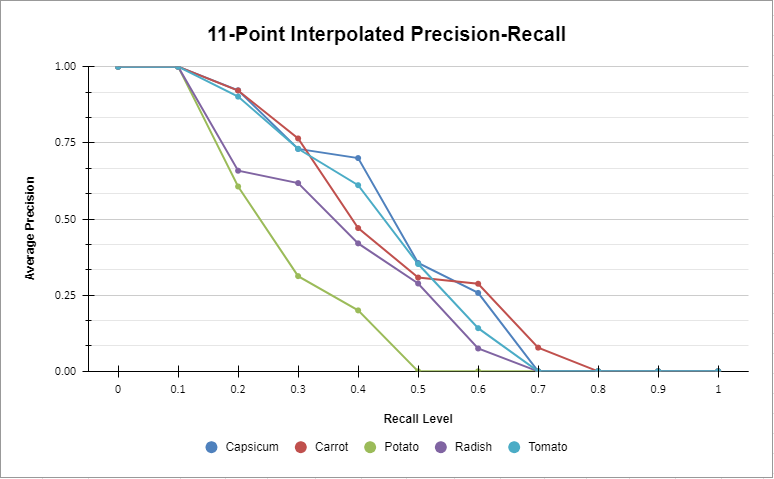
The Google Sheet link: <https://docs.google.com/spreadsheets/d/1NvmvlkUVbhdynONyJ4IbvP9H-fwjOYP7/edit?usp=sharing&ouid=115005893896843335760&rtpof=true&sd=true>

In the line graph shown below, we have displayed the results of our Precision-Recall calculation for the project. This graph presents a comprehensive overview of the precision and recall values obtained for the whole Content Based Image Retrieval system. By analyzing this data, we found out that from 0 to 0.1 recall level, the system achieves a high precision value, indicating that it accurately classifies most positive images. However, this comes at the expense of lower recall, suggesting that some positive images may have been missed by the system.



The system fails to maintain higher precision throughout the process as the average precision falls straight from 1.00 to 0.48 from recall level 0.1 to 0.4. This is because the similarity measures are retrieving images with higher False Positive results. The system also has average precision of 0.01 at recall level 0.7 which indicates that almost all of the results retrieved by the system are going to be False Positive reflecting the extremely low precision shown.

When using the HSV color histogram in CBIR, as the system aims to retrieve more images, it may become less selective in its classifications and include images that are not truly similar. This can lead to an increase in False Positives, where images that are not actually similar are considered as matches. The falling precision as recall increases suggest that the system is becoming more inclusive in considering images are similar, even though they may have significant differences in their color distributions. This could be due to the nature of the color histogram representation, where slight variations in color can lead to similar histogram values, resulting in False Positive matches.



From the graph, we observed that capsicum’s query has the most consistency of performance followed by radish, carrot, tomato and potato. This is due to capsicum’s query having the most relevant document retrieved than others which results in more specific and accurate results. Moreover, we can conclude that capsicum’s images have distinct visual features such as color, shape and size characteristics that are well-captured by the HSV colour-histogram features, resulting in more consistent and accurate retrieval results. Consistency in the visual characteristics of capsicum images reduces the chances of confusion or ambiguity during the retrieval process, leading to more consistent and accurate results. Other than that, the Minkowski distance metric we used as the similarity measure might effectively capture the dissimilarities between capsicum images and other vegetable images. If capsicum images are well-separated from other vegetable images in the feature space, the retrieval system can successfully identify and retrieve capsicum images with higher precision.  
  
Meanwhile, potato’s query has the worst retrieval among the vegetables. This is because histograms may not be the most effective feature representation for capturing the visual characteristics of potatoes. Potatoes can have relatively uniform colors and textures, which might lead to similar histograms across different potato images. Consequently, the lack of distinctive features can make it more challenging for the retrieval system to accurately distinguish between potato images, resulting in lower retrieval performance. Other than that, the retrieval performance depends on the visual similarity between the query potato image and the images in the database. If the potato image shares visual characteristics with other vegetables or if there are visual similarities between different vegetable images, it can lead to less accurate retrieval results.

**Project Reflection**

| Name | Reflection |
| --- | --- |
| Amir Nurhakim Bin Mohd Zaid | From this project, I can see the back-end of multimedia retrieval that is happening on the internet. I also somewhat gain more knowledge on how to work around Excel with the formula and graphs. I am also very thankful to have these teammates as the teamwork is very immaculate. |
| Anas Zulkifli bin Mohd Jeffry | I learn a lot of new things especially about content-based image retrieval. I feel that the thing behind the information retrieval is not so simple, I need to do many processes like pre-processing, similarity measure and evaluation. My suggestion to improve this course is do the programming practice to implement the retrieval system |
| Nik Muhammad Asyraf bin Nik Ismail | Working on this project has been an exciting and valuable experience for me. It provided me with a hands-on opportunity to apply the knowledge and skills I acquired throughout the course. |
| Muhammad Amirul Afiq Bin Juma’at | From this assignment, I learned how good teamwork made the work easier. I also learned how to use Excel sheets. I also experienced roughly how MMIR is done in the framework. |

**References**

Ahmed, M. (2021). *Vegetable Image Dataset* [Data set]. Kaggle.

<https://www.kaggle.com/datasets/misrakahmed/vegetable-image-dataset>.

Cao, X., Liu, Y., Yang, Y., & Huang, T. (2017). *Deep neural networks for cross-modal multimedia retrieval: A survey.* ACM Transactions on Multimedia Computing, Communications, and Applications, 13(1s), 1-19.

Marshad, B. (2021). *CBIR--Content-Based-Image-Retrieval* [GitHub repository]. GitHub.

<https://github.com/bassamarshad/CBIR--Content-Based-Image-Retrieval>.

Wang, S., Huang, Z., & You, X. (2018). *Deep hashing for large-scale image retrieval: A survey.* IEEE Transactions on Multimedia, 20(10), 2564-2578.

Yu, L., & Liu, X. (2019). *Deep image retrieval: A survey.* Journal of Visual Communication and Image Representation, 63, 102535.

Yu, F., Ji, R., Liu, Y., & Chen, J. (2017). *Learning a deep compact image representation for visual tracking.* IEEE Transactions on Image Processing, 26(7), 3491-3504.

Zhang, L., Liu, W., Wu, J., & Ji, R. (2018). *Learning deep binary descriptor with multi-quantization for fast image retrieval*. IEEE Transactions on Image Processing, 27(1), 210-223.