1. Experiment process

   Dataset:

   MRFLICKR is an open evaluation project with 25,000 images downloaded from the social photography site Flickr through its public API with complete manual annotations. The dataset's annotations for each image provide some categories of what is inside the image but not as accurate and sometimes too broad.

   Process:

   Randomly query an image from the flicker dataset. Query the image in the system and check 4 categories mentioned below. Choose a single ROI within the image and query the image again with Region-Of-Interest. Check the retrieved images and record observations. The system has been modified to take multiple ROI later but this experiment was done with single ROI region.

   Dictionary:

   The system is refers to the image retrieval service currently available on Traffic Jam. Images are one single image chosen as the query image that will be fed into the system for similar image searching. Region-of-Interest is referred to as ROI which is a selected region within the image being queried. The system behaves or works well when the retrieved/result images contain highly relevant or similar/same objects that exists in the query image.

2. 4 Categories

   Pattern:

   Pattern that exists within the entire image or the ROI region. For example, if there are grapes in the image, then the system will retrieve round objects that are gathered together like candies, round fruits, and pebbles. Also, if there is a knitted pattern in the image, the system will attempt to find images with the pattern.

   Color:

   Colors that exists within the entire image or the ROI region. For example, if there is a specific color that stands out in the image like red, the retrieved images have objects that have red color in them.

   Shape:

   Refers to the shape of the objects or human poses in the image or the ROI region that is identifiable by the system. Also, specific ROI chosen area acts as shape depending on which area or what object shape has been chosen as the ROI. If a chair is in the middle of the image, the retrieved images have furniture with legs or similar shaped objects. Human's face is considered as a round shape and laptop or computer screens take a

square shape in the images.

Structure:

Refers to the entire structure of the image with or without ROI region. If the ROI region is too big, then it divides the image and the system seems to recognize that as a structure within the entire image. For example, if there is a scenery image where top half of the image is clouds and bottom half is mountains, then the retrieved images will have that structure as well. If a wide area like top half of the image is chosen as ROI, it tends to divide the image there and retrieves images that are divided in that way. Also, when a stuffed animal in the middle of the image as chosen as ROI, the system retrieves images that have objects that have similar color in the middle of the image and mostly not stuffed animals.

Out of all 4 categories, pattern and color influences the image retrieval the most when they exist in the image. However, if there is no eye-catching pattern or color in the image, the system retrieves images that have similar shaped object images or similar structured images.

3. General issues observed

When the ROI region is too small, then the system does not recognize the object and retrieves blurry images. Some objects contextually matter and when removed from its environment, chosen as ROI, the system seems to get confused. This happens a lot in scenery images.

When ROI is chosen, if the system does not recognize it, it tends to blur the shape and retrieve images with similar structure with objects of the blurred shape in ROI region. For example, when unidentifiable objects are chosen in the image as ROI, the retrieved images have shadows or pillars in the ROI region. Images with shadows of objects/subjects are picked up in these situations.

Also, if the system was not trained on a certain object, or the dataset does not have the image that contains the same object, it does not seem to work well. The system recognizes well with subjects/objects that was trained on but not strong when it comes to new subjects/objects.

4. Observations by case

Area of ROI:

Area of the ROI was chosen manually considering the labels available. Most of the time, the main object visible in the image was chosen. Exceptions were the scenery images where the entire sky or the mountain area was chosen. Though choosing a wide and non-shaped area of the image does not work well.

Objects:

Refers to any indoor and outdoor objects that exists in the image. Typical objects include furniture, stuffed animals, laptops, food, books, and more. Exact object retrieval is not available but shape and color is well picked by the system.

Transportation label exists for the images from MIRFLICKR database but does not recognize well in the system.

Human:

Refers to any image that has human in it. Regardless of how visible or apparent the face is. The system recognizes faces well but does not differentiate age or gender.

Portrait:

Refers to one person, once face, one upper body images. The system retrieves portraits very well and similar structures. Gender recognition is poor but it recognizes faces pretty well without the ROI. However, once face is chosen as ROI, shape and structure plays some role and picks up moon photos along with face photos.

Group:

Refers to multiple people, face, upper body images. The system recognizes well that there are multiple people in the image and retrieves images with multiple people or standing figures.

Scenery:

Refers to landscape and outdoor scenery with specific structure and color. The system tends to retrieve scenery well.

Examples

-mountain and sky/clouds combination

-mountain and reflection in the water (normally lakes or rivers)

-photos near water

-grass field

Animals:

Refers to any image that contains animals that are main focus of the image. Images with flock of birds or shadows are not applicable. Frequent examples were dogs, cats, horses, and birds. Dogs and cats were well identified by the system. Dogs seem to be better identified from this database than the cats. When cat images were queried as ROI, the system seems to pick up animals with pointy ears were also retrieved. Birds are recognized as well most of the time. However, when ROI is used for the bird itself, some images with flowers are retrieved due to strong pattern and color in the feathers.

Food:

The system seems to recognize when food is in the image but does not know what type it is. When food image is queried, other food images are retrieved. However, this may e happening due to lack of specific type food in the database. For example, if grapes are in the image and chosen as ROI, round snacks and objects were retrieved.

Time of day:

Using color, the system recognizes day/night pretty well especially in scenery photos. However, it seems to pick up day photos better than the night photos. Using ROI with night photos work less prominently than choosing a ROI from a day photo. For example, choosing a bridge from a city night view does not work well.

5. Suggestions and questions for ROI implementation

-Losing context: category box for ROI to let the system know the category.

-Overpowering color: color ignore button.

-Recommend using feedback retrieval.

-Train system depending on the need of the retrieval.

-Choosing a specific type database to train the system and then use that to query that type of images is likely to be more effective with or without ROI

6. Summary

Overall, image retrieval system does a good job in identifying human, animals, and scenery images which the system was trained on. However, it struggles identifying objects in the image that are not trained on or if the dataset does not contain enough of the same object. Similar patterns, colors, shapes, and structures are strongly observed between the query image and the results. Except for few objects that matter within a context, choosing a ROI to retrieve images seems to do a better job in finding images with similar patterns, colors, shapes, and structures, not necessarily the same object.