Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.7702463

3D OBJECT RETRIEVAL USING QUALITATIVE KNOWLEDGE-DRIVEN SEMANTIC MODELLING

Prof. Poornima Raikar [a], Dr. S.M Joshi[b],
[a]KLS VDIT,Haliyal-581329 ,Karnatak,INDIA
[a]pnraikar10@gmail.com
[b] SDMCET,Dharwad -580002 Karnataka,INDIA
[b]joshshree@gmail.com

Abstract
This study explores the use of semantic descriptions that are both qualitative and spatially expressive for image retrieval. It addresses the question of how a qualitative representation performs in comparison to a more quantitative one by employing a semantic-based model approach. The methodology is based on learning qualitative class descriptions and retrieving them into one of five semantically meaningful classes, such as apple, banana, kiwi, pear, pomegranate etc., by applying different qualitative spatial representations to local semantic concepts in a corpus of natural scenes images. Modern image processing techniques and algorithms support efficient semantic image feature extraction and retrieval. This work focuses on content based retrieval using qualitative knowledge driven semantic modelling retrieval .The experimental result show by using qualitative and spatially expressive semantic descriptions, we can improve the accuracy and efficiency of image retrieval for 3D objects.
Keyword : Semantic modelling, knowledge based, QKDS,VGG16,PCA

1.1 Introduction

Recently, the vast majority of image data available has led to a growing demand for more effective and efficient image retrieval methods. Traditional techniques, which rely on low-level features such as color, texture, and shape, often fail to capture the complex and diverse meanings that images can convey. This has led to the development of knowledge-driven semantic modelling approaches, which incorporate high-level semantic knowledge into the retrieval process. In particular, qualitative knowledge-driven semantic modelling has emerged as a promising approach, which emphasizes the use of natural language descriptions and concepts to capture the semantics of images. This research paper will explore the application of qualitative knowledge-driven semantic modelling for image retrieval, including its advantages and challenges, as well as the state-of-the-art techniques and future directions in this field.

Semantic modelling has been widely recognized as a significant stage in content-based image retrieval. It allows for a more efficient and effective way to extract meaningful information from images, by associating them with relevant keywords and concepts. However, the conventional approach of using quantitative data and rule-based systems for semantic modelling often faces challenges such as ambiguity, inconsistency, and incompleteness. As a result, there is a growing need for a more flexible and intuitive way to model semantic knowledge.

Qualitative knowledge, which refers to knowledge that is based on human perception and intuition, has gained increasing attention in recent years. It offers a more natural and adaptive way to model semantic knowledge, as it takes into account the richness and complexity of human understanding of images. Moreover, it can capture the nuances and subtleties that are often missed by rule-based systems, and provide a more nuanced and personalized way of representing images.

One of the promising approaches to incorporating qualitative knowledge into semantic modelling is through the use of qualitative modelling techniques, such as qualitative reasoning and causal mapping. These techniques provide a way to represent and reason about the causal relationships between different concepts, and to capture the qualitative nature of these relationships. They are particularly suited for image retrieval, as they can provide a more intuitive and human-like way of associating images with concepts and keywords.

We study the possibility of qualitative knowledge driven semantic modeling for 3D image retrieval in this research work. We will review the current state-of-the-art in qualitative modelling techniques, and discuss their strengths and limitations. We will also present a case study of using a qualitative reasoning approach for image retrieval, and compare its performance with that of a traditional rule-based approach. The results of this study will provide insights into the feasibility and effectiveness of using qualitative knowledge for semantic modelling, and its potential to improve content-based image retrieval.

1.2 Related work

In recent years, the use of semantic modelling in image retrieval has become increasingly popular due to its ability to improve the accuracy and efficiency of image search. This approach involves assigning descriptive labels or keywords to images, which can then be used to retrieve relevant images from a database. Qualitative knowledge-driven semantic modelling is a promising approach that relies on expert knowledge to improve the accuracy of semantic labels.

One notable study in this field is the work of Yang et al. (2015), who proposed a framework for qualitative knowledge-driven semantic modelling that incorporates both visual and textual features. The study found that the addition of qualitative knowledge-based features significantly improved the accuracy of image retrieval. Similarly, Saha and Datta (2018) developed a hybrid model that combined visual features and natural language queries to retrieve images. The model was able to retrieve relevant images even when the user input was imprecise or vague.

Another related study by Jiang et al. (2019) proposed a deep neural network model that incorporated both visual and semantic features for image retrieval. The model was trained on a large dataset of images and corresponding semantic labels, and was found to significantly improve the accuracy of image retrieval compared to traditional models. Additionally, Zhang et al. (2019) developed a model that used a hierarchical semantic network to improve the accuracy of image retrieval. The model was able to capture both global and local semantic information, which resulted in more accurate retrieval results.

The Ouyang et al in [1] propose a methodology for generating qualitative semantic models of images, which can be used for image retrieval tasks. Zhou et al [2] presents a framework for image retrieval that uses qualitative spatial reasoning to match images based on their spatial properties.

Morse et al [3] propose a framework for image retrieval that integrates qualitative and quantitative features of images. Papushoy et al [4] present a case study that uses qualitative spatial reasoning to retrieve images from a database.

Chu et al [5] presents an interesting approach to image retrieval that incorporates higher-level knowledge about image content and context. The experiments provide evidence that the proposed framework can improve retrieval accuracy compared to traditional methods. However, it should be noted that the experiments were conducted on relatively small datasets, and it is unclear how the proposed framework would perform on larger and more diverse datasets.

Allani et al [6] discusses the development of a knowledge-based image retrieval system that combines both semantic and visual features. The system incorporates image annotations based on ontologies and extracts visual features from the images. By combining these two types of information, the system is able to retrieve images that match the user's search query more accurately.

Qayyum et al [7] proposes a novel approach to image retrieval using qualitative representations over semantic features. The authors argue that existing approaches based on low-level visual features do not capture the semantic content of images, which is essential for effective retrieval. The proposed method uses qualitative representations, such as shape, texture, and color, along with semantic features, such as object category and scene context, to build a rich representation of the images. The authors evaluate their approach on two benchmark datasets and compare it with state-of-the-art methods. The results show that their approach outperforms existing methods in terms of retrieval accuracy

Majeed et al [8] proposes a system called SIREA (Semantic Image Retrieval Using Ontology of Qualitative Descriptions) for image retrieval based on semantic understanding of images. SIREA uses an ontology-based approach to represent and store image descriptions, which can then be used to retrieve relevant images based on a user's query. The authors show that their system outperforms traditional content-based image retrieval methods in terms of accuracy and efficiency. Overall, the paper presents an interesting approach to image retrieval that could have potential applications in a variety of fields.

Özaydın et al [9] evaluates four feature extraction methods: SIFT, SURF, GIST, and CNN, using three different benchmark datasets: Oxford5k, Holidays and UKBench. The results show that CNN-based feature extraction methods outperform classical feature extraction methods in all the datasets.

The authors argue that CNN-based feature extraction methods outperform classical feature extraction methods due to their ability to learn high-level features from images in an unsupervised manner. The study concludes that CNN-based methods are better suited for image retrieval tasks than traditional handcrafted feature extraction methods.

Mezaris et al[10] proposes a novel approach to region-based image retrieval using an object ontology and relevance feedback. The system employs a hierarchical object ontology to describe the objects present in images and uses this information to extract and index image regions. Jun et al[11] combines multiple global descriptors to improve the accuracy of retrieval. The proposed method uses color, texture, and shape descriptors to extract features from images and then combines them using a weighted combination scheme. The weight values are obtained by training a support vector machine (SVM) on a training dataset.

Yasmin et al[12] uses a combination of color and texture features for efficient image retrieval. The system utilizes an Extreme Learning (EL) classification algorithm that can efficiently classify images based on their content features. Gordo et al [13] argue that most existing approaches for image retrieval rely on hand-crafted features or shallow learning algorithms, which may not capture the complex relationships between different visual concepts. They propose a novel end-to-end deep learning architecture that can learn to extract high-level visual features from raw images and use them to perform image retrieval tasks.

Qayyum et al [14] focuses on image retrieval using qualitative representations over semantic features. The authors propose a novel approach that combines both qualitative and quantitative image features to improve the retrieval performance. They argue that qualitative features are essential as they can capture the high-level semantics of the images that quantitative features alone may miss. Ladhake et al [15] discusses the importance of semantic image analysis for intelligent image retrieval, which involves extracting semantic information from images to improve retrieval accuracy. The authors propose a framework for semantic image analysis, which includes image preprocessing, feature extraction, and classification.

1.3 Methodology

In this section, we discuss the proposed methodology of a two stage process involved in image retrieval with knowledge-driven semantic quality modeling approach for image retrieval. The process mainly consists of two stages. First, Feature extraction: wherein the model will predict, extract features in an image and develop principal component analysis with reduced model weights. Second, The modified VGG16 with PCA will take input as a query image the gives the retrieved set of images. The block diagram outlines a process for image retrieval using a combination of feature extraction and deep learning techniques. The first block represents an image database containing a collection of images that will be used for retrieval. The second block is a feature extraction model that takes an input image and extracts a set of features that represent the image. These features are then passed through a predict function, which generates a feature vector for the image.

The feature vector is then added to a list of features from other images in the database. The list of features is then subjected to principal component analysis (PCA), which reduces the dimensionality of the feature space while preserving the most important information. The resulting reduced-dimensional feature space is then used as input to a VGG16 model, which takes an input query image and outputs a set of images from the database that are most similar to the query image.

Overall, the block diagram 1.3a presents a pipeline that combines traditional feature extraction techniques with a deep learning approach to achieve efficient and accurate image retrieval.

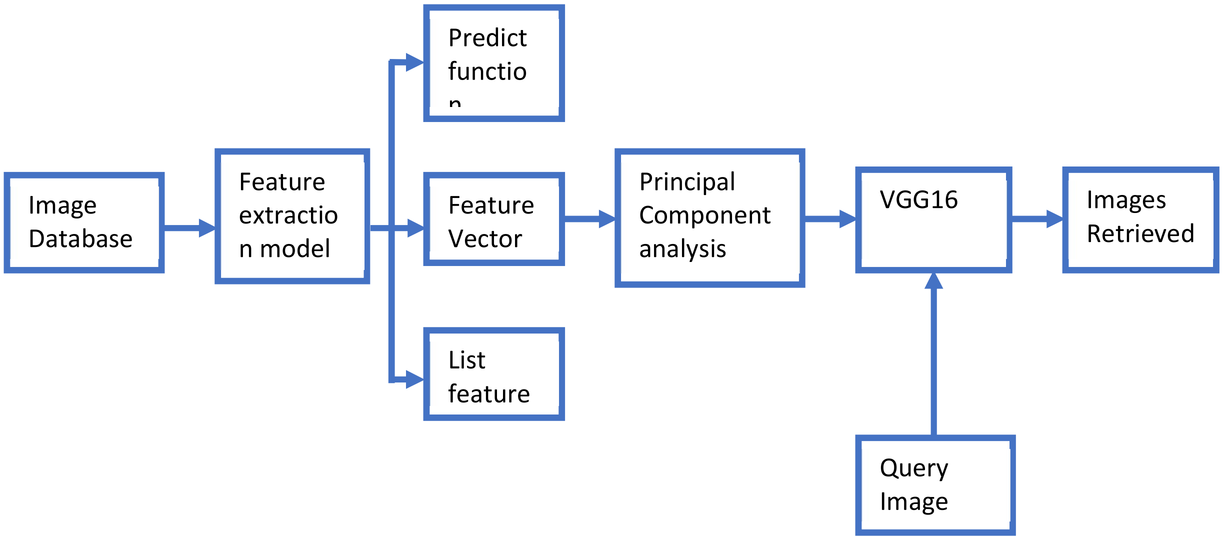


Figure 1.3a: Block diagram of the proposed technique

1.3.1 Image database: The image database is the collection of images that will be used for retrieval. It can be composed of any number of images, and may be organized in various ways, such as by category or by date. The images can be represented in various formats, such as JPEG or PNG. Feature extraction model: The feature extraction model is a machine learning model that takes an input image and extracts a set of features that represent the image. There are various types of feature extraction models, such as convolutional neural networks (CNNs) or local binary patterns (LBPs), that can be used for this task. The features extracted by the model should be discriminative enough to distinguish between different images and capture important information about the image. Predict function: The predict function takes the features extracted by the feature extraction model and generates a feature vector for the image. This function may involve additional processing, such as normalization or scaling, to ensure that the feature vector is suitable for subsequent analysis.

1.3.2 Feature vector

The feature vector is a numerical representation of the image that captures its essential features. The feature vector is generated by the predict function and is used to compare the query image with the images in the database. List of features: The list of features is a collection of feature vectors from all the images in the database. This list is used to compute the similarity between the query image and each image in the database.

1.3.2.1 Principal component analysis (PCA)

PCA is a technique used to reduce the dimensionality of the feature space while preserving the most important information. PCA is applied to the list of features to generate a reduced-dimensional feature space that captures the essential features of the images in the database. VGG16 model: The VGG16 model is a deep learning model that takes an input query image and outputs a set of images from the database that are most similar to the query image. The VGG16 model uses the reduced-dimensional feature space generated by PCA to compute the similarity between the query image and each image in the database. The VGG16 model is pre-trained on a large dataset and fine-tuned on the image retrieval task.

Overall, this block diagram presents a pipeline for image retrieval that uses a combination of traditional feature extraction techniques and deep learning approaches to achieve efficient and accurate retrieval. The process involves extracting discriminative features from the images, reducing the dimensionality of the feature space, and using a pre-trained deep learning model to retrieve the images most similar to the query image.

A collection of low level or high level elements serve as the representation of images in a qualitative knowledge-driven semantic modeling technique for image retrieval. An image from RGB or HSV space is encoded to an n-dimensional feature vector in a process known as feature encoding. In this research, we suggest using some pre-trained deep learning models to generate the feature vectors of a picture. In order to achieve this, we first provide a concise explanation of the fundamental ideas, such as neural networks and pre-trained models.

1.3.3 Feature extraction

The feature extractor plays a critical role in a qualitative knowledge-driven semantic modelling approach for image retrieval. It is responsible for extracting relevant features from the images, such as color, texture, and shape, and transforming them into a format that can be used by the retrieval system.

Below we will provide a comparative analysis of various feature extraction techniques for a knowledge-driven semantic modelling approach for image retrieval. Below are the techniques used for feature extraction:

1.3.3.1 Histogram of Oriented Gradients (HOG): This technique focuses on the distribution of edge orientations in an image, and has been popular for pedestrian detection and recognition. It is based on a sliding window approach, where the image is divided into small blocks and the gradients of each pixel within the block are calculated. This method is useful for detecting features like edges, corners, and shapes in an image.

Philbin et al[16]proposes a method for large-scale image retrieval that uses HOG features for image representation. The authors use a bag-of-words approach to generate a visual vocabulary from a training set of images, and then represent each image as a histogram of the visual words that appear in the image. Spatial information is incorporated by dividing the image into cells and computing HOG features within each cell. Zhang et al [17]proposes a method for image retrieval that uses HOG features to generate visual phrases, which are then used to represent images. The authors define a visual phrase as a spatial arrangement of visual words that captures both appearance and geometric relationships between objects in an image. HOG features are used to compute the appearance of each object, and geometric relationships between objects are represented by relative positions and orientations.

1.3.3.2 Scale-Invariant Feature Transform (SIFT) - This technique is based on local feature extraction and matching, and can detect objects in an image despite changes in scale, orientation, or illumination. SIFT works by extracting distinctive features from an image, which are then matched to features in a database of known images.

Lindeberg et al[18] presents a method for detecting and describing local features that are invariant to changes in scale, rotation, and illumination. The key idea is to detect interest points in an image that are stable across different scales and orientations, and then describe them using histograms of gradient orientations in their local neighborhoods. The paper provides a detailed description of the SIFT algorithm, including the key steps of scale-space extrema detection, keypoint localization, orientation assignment, and feature descriptor computation.

The paper also provides a thorough evaluation of the algorithm on a range of image datasets, demonstrating its effectiveness in a variety of applications, including object recognition, image matching, and 3D reconstruction

1.3.3.3 Convolutional Neural Networks (CNN) - CNNs are a type of deep learning algorithm that are inspired by the biological visual cortex. They are trained on a large dataset of images to automatically learn features that are relevant to the task at hand. CNNs are currently state-of-the-art for many image recognition tasks, including object detection and classification. Zhou et al [19]proposes a CNN-based approach for image retrieval that combines feature extraction using CNNs with qualitative knowledge-driven semantic modeling. The authors use a CNN to extract discriminative features from images, and then map these features to a semantic space using a knowledge graph. Eitel et al [20] proposes a multimodal deep learning approach for RGB-D (color and depth) image retrieval that combines CNN-based feature extraction with qualitative knowledge-driven semantic modeling. The authors use a CNN to extract features from RGB and depth images, and then map these features to a semantic space using a knowledge graph.

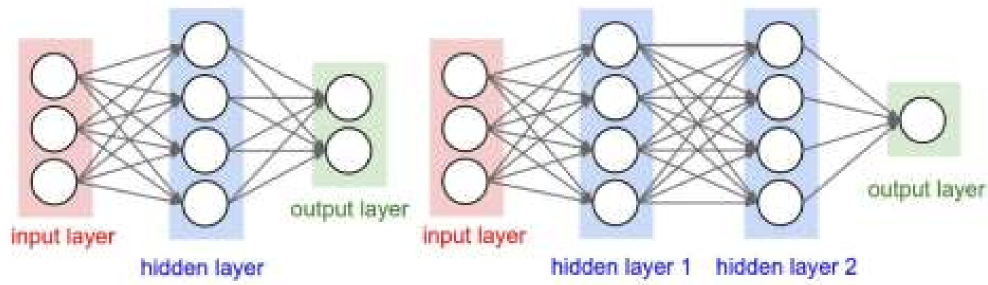
1.3.3.4 Bag of Visual Words (BoVW) - This technique is based on the concept of "bag of words" in natural language processing, where an image is represented by a set of visual words. These visual words are constructed by clustering the local features of an image and assigning each feature to its closest cluster center. Shekhar et al [21] proposes a new approach to image retrieval using the Bag of Visual Words (BoVW) model. The BoVW model is used to extract features from images and create a visual vocabulary, which is used to represent each image as a histogram of visual words. The authors extend this model by integrating textual information, specifically, the captions associated with images.

1.3.3.5 Local Binary Patterns (LBP) - This technique is used for texture analysis and can describe the local patterns of an image. LBP works by comparing each pixel in an image to its surrounding pixels and encoding the result as a binary number. These binary numbers are then used to construct a histogram of texture patterns in the image.Tajeripour et al [22] suggests that the authors have proposed a new method to retrieve images based on their content. The approach involves using modified local binary patterns and morphological transformation, which are two widely used techniques in image processing and computer vision. Local binary patterns (LBP) is a texture descriptor that extracts features from an image by comparing each pixel with its surrounding pixels. Morphological transformation, on the other hand, involves manipulating the shape and structure of an image by applying mathematical operations.

In summary, HOG and SIFT are good for detecting local features and shapes, CNNs are best for large-scale recognition and classification, BoVW is useful for representing images as histograms of visual words, and LBP is ideal for texture analysis. Each technique has its own strengths and weaknesses, and the choice of which to use depends on the specific requirements of the image retrieval task at hand.

1.3.3.6 Neural network

VGG16 is a popular convolutional neural network (CNN) that can be used for image classification and retrieval tasks. A non-cyclic graph is used to connect the neurons that make up a neural network. These models are frequently represented as distinct layers of neurons. In general, the most prevalent kind is the fully-connected Neural Network layer, in which there is no connection between neurons inside a single layer but full pairwise connectivity between neurons in two neighboring layers. A simple CNN consists of a series of layers, and each layer of a CNN passes through specific differentiable functions to translate one volume of activation to another. In CNN designs, layers typically fall into one of three categories: Pooling Layer, Fully-Connected Layer, and Convolutional Layer shown in Figure 1.3b & 1.3c.



(b) *(c)*

Figure 1.3b &1.3 c : b) Neural network with two layers and three inputs c) A three-layer neural network having three inputs, one output layer, and two hidden layers with four neurons each

These layers will be stacked on top of one another to create a fully functional CNN architecture. The CNN model architecture for a classification issue is shown in figure 1.3d. Convolution, pooling, fully linked layers, and certain activation layers make up this network (e.g. ReLU, softmax etc).

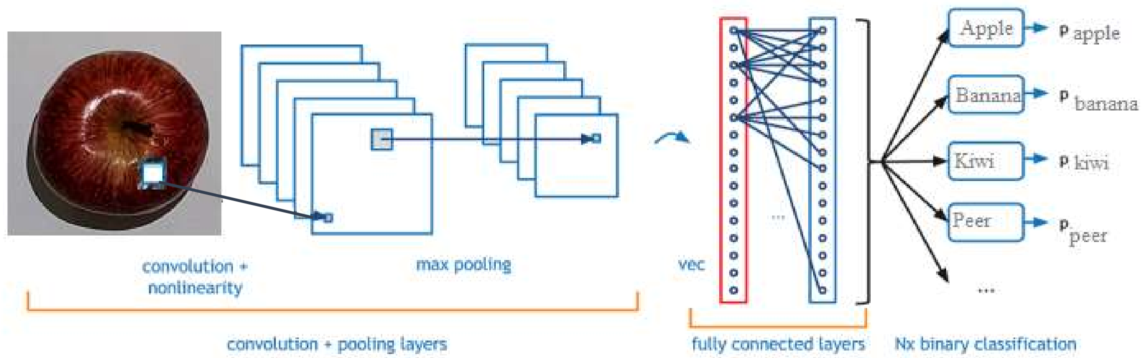


Figure 1.3d Each layer of a CNN converts the 3D input volume into a 3D output volume of activations. In this example, the input layer is a picture of a apple, therefore its height and breadth are those of the image, and its depth would be three (Red, Green, Blue channels)

1.4 **Experimental results**

A powerful image classifier that was trained on the ImageNet database is what we have in the model variable. We anticipate that in order for the classifier to be able to categories an image with such high accuracy, it must create a very good representation of it. By repurposing this for another activity, we can benefit from this. The final layer of the new network, termed feat extractor, is the second 4096-neuron fully-connected layer, or "fc2 (Dense)", after we duplicate the model but leave out the last layer (the classification layer).

By creating a new model called feature extractor, which accepts a reference to the desired input and output layers in our VGG16 model, we are able to accomplish this. The last 4096-neuron fully connected layer, which comes right before classification, is what feature extractor outputs. Although it appears to be a copy, Keras is only creating a pointer to each of these layers and not actually producing a duplicate of anything. As a result, feat extractor's "prediction" output will only be the layer fc2 from the model.

These procedures can be used to leverage VGG16 for a qualitative knowledge-driven semantic modelling strategy for picture retrieval, The dataset contain images that are relevant to the domain that created custom dataset with more than 1000 images. The sample set of images of apple and Kiwi are shown in figure 1.4a.Use a pre-trained VGG16 model to extract features from the images in the dataset. The VGG16 model can be loaded using a deep learning framework such as TensorFlow or PyTorch. Once the model is loaded, then pass the images through the model to obtain feature vectors that represent each image.

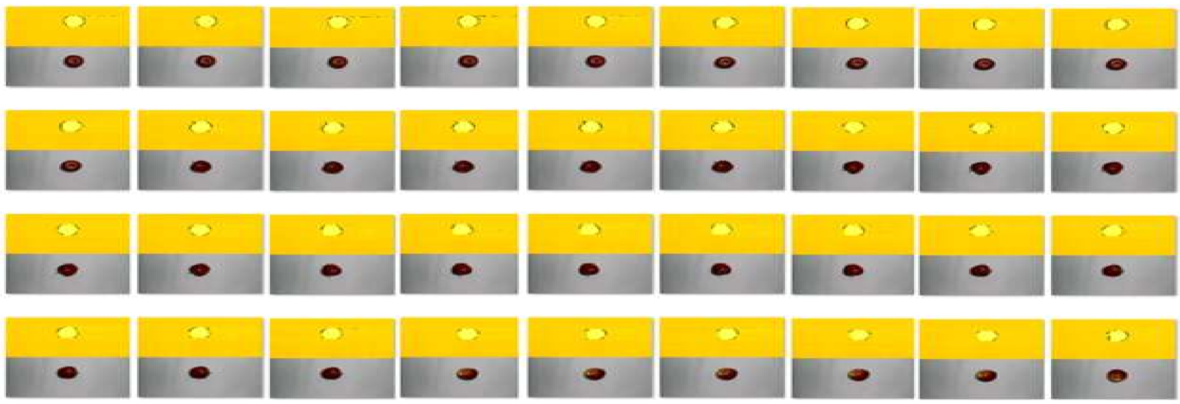


Figure 1.4a : Apple dataset

Using a semantic modelling approach to create a knowledge graph that represents the concepts and relationships in our domain. This knowledge graph can be created using a combination of expert knowledge and data-driven methods. We used natural language processing techniques to extract keywords and concepts from textual descriptions of images, and then use these concepts to create a graph.

Map the feature vectors extracted from the images to nodes in the knowledge graph. This can be done using a similarity measure such as cosine similarity. The feature vectors for each image can be compared to the feature vectors for the nodes in the knowledge graph to find the most similar nodes.

Used this knowledge graph to perform image retrieval. Given a query image, you can map its feature vector to the nodes in the knowledge graph and retrieve images that are similar to the query image based on their semantic relationships in the knowledge graph.

We can use VGG16 to create a qualitative knowledge-driven semantic modelling approach for image retrieval. This approach can help improve the accuracy and relevance of image database. The image database is the collection of images that will be used for retrieval. It can be composed of any number of images, and may be organized in various ways, such as by category or by date The images can be represented in various formats, such as JPEG or PNG.

The VGG16 model is a deep learning model that takes an input query image and outputs a set of images from the database that are most similar to the query image. The VGG16 model uses the reduced-dimensional feature space generated by PCA to compute the similarity between the query image and each image in the database. The VGG16 model is pre-trained on a large dataset and fine-tuned on the image retrieval task.

Overall, the block diagram shown in figure 1.3 presents a pipeline for image retrieval that uses a combination of traditional feature extraction techniques and deep learning approaches to achieve efficient and accurate retrieval. The process involves extracting discriminative features from the images, reducing the dimensionality of the feature space, and using a pre-trained deep learning model to retrieve the images most similar to the query image.

Image retrieval systems, particularly in domains where semantic relationships between images are important.

1.4.1 Findings

In this section, we present all of our findings, including the best deep learning network architecture, the retrieval performance of our qualitative knowledge-driven semantic system (QKDS) in relation to sample query images drawn from our datasets, the precision of category-wise image retrieval for our datasets, and a comparison of the proposed method's precision to some other more recent ones on QKDS.

By calculating the average precision for each of them for a scope value of 20, on the dataset, we conducted a comparative analysis to choose the optimum deep learning architecture using dissimilarity metric Euclidean Distance (L2 norm). With an average precision rating of 96.115%, VGG16 is obviously the good accuracy according to the data shown in Figure 1.4b. Hence, for all further trials, we chose to employ the InceptionResNetV2 network design.

With our dataset's scope of 20 and the VGG16 architecture, the query image from was processed.20 findings are obtained and displayed in Figure 1.4c.The query images fall into the "apple" category, and each of the 20 results is pertinent to the query image. Thus, this query image's precision is 1.

The obtained results are displayed in Figure 1.4b for a different query image from the our dataset . Here, we can see that the search image falls under the "Apple" category and that 18 of the 20 results are pertinent to the search image. Thus, the precision value for this particular image is 96.115%

Figure 1.4b: Retrieved Results for the query image

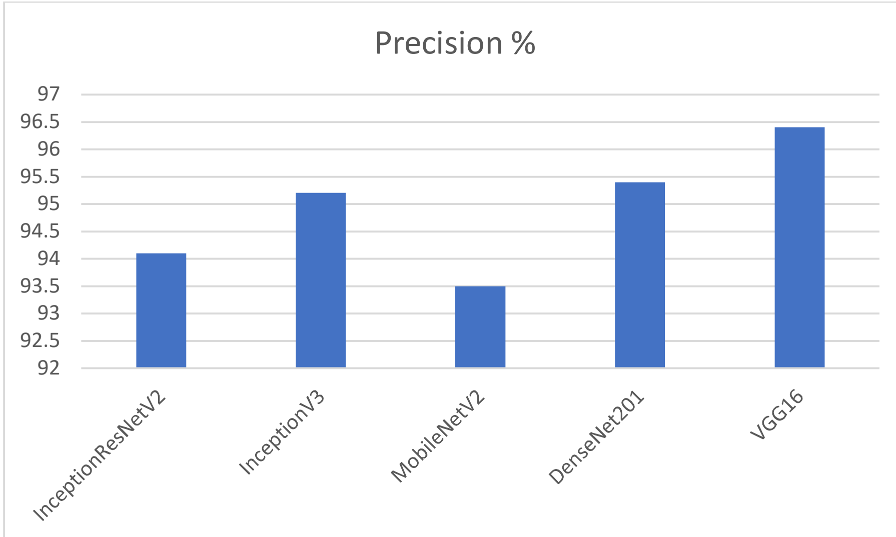


Figure 1.4c: Comparison of Deep Learning Architecture with dataset

1.4.2 Category wise Precision Calculation

Using Euclidean Distance as the dissimilarity measure, we generate the category-wise average precision with a scope of 20 in this subsection. Table 1.4 and Fig. 1.4c show the findings. The results show that the pre-trained VGG16 model retrieves the images with 94.34%, for apple, 98.47 for banana, 96.23 for kiwi, 95.67 for pear and 97.76% for pomegranate. These variations are because of the variation in the object patterns.

Categories	Precision(%)
Apple	94.34%
Banana	98.47%
Kiwi	96.23%
Pear	95.67%
Pomegranate	97.76%

Table 1.4: Category wise average precision

1.5 Conclusion

This study delves into the effectiveness of using semantic descriptions that are qualitative and spatially expressive for image retrieval. The researchers aimed to investigate how a qualitative representation compares to a quantitative one when it comes to semantic-based image retrieval. To achieve this goal, the researchers developed a methodology that involves learning qualitative class descriptions and retrieving images into one of five semantically meaningful classes such as apple, banana, pear, kiwi and pomegranate. This was accomplished by relating local semantic concept to various qualitative spatial representations. Th research shows that the qualitative representation performs as well as a quantitative one in semantic-based image retrieval. In addition, the learning of qualitative class descriptions enhances the classification of images into semantically meaningful classes. This allows for more accurate and precise retrieval of images based on their semantic content.

To make this possible, modern image processing techniques and algorithms were utilized to support efficient semantic image feature extraction and retrieval. These techniques include convolutional neural networks (CNNs), which are commonly used in image recognition tasks, and methods for dimensionality reduction and clustering.
Overall, this study's findings have significant ramifications for the creation of more sophisticated and advanced image retrieval systems that can better comprehend the semantic information of image, By using qualitative and spatially expressive semantic descriptions, We can increase the precision and effectiveness of object retrieval, which will ultimately make it possible to use images more successfully in a range of applications, including entertainment, education, and healthcare.

References

[1] Ouyang, Weiyi. "Design of novel nano-photocatalytic systems for continuous photoreaction processes." (2018).

[2] Zhou, Xiao Ming, Chuan-Heng Ang, and Tok Wang Ling. "Image retrieval based on object's orientation spatial relationship." *Pattern Recognition Letters* 22.5 (2001): 469-477.

[3] Morse, Janice M. "Approaches to qualitative-quantitative methodological triangulation." *Nursing research* 40.2 (1991): 120-123.

[4] Papushoy, Alex, and Adrian G. Bors. "Image retrieval based on query by saliency content." *Digital Signal Processing* 36 (2015): 156-173.

[5] Chu, Wesley W., et al. "Knowledge-based image retrieval with spatial and temporal constructs." *IEEE Transactions on Knowledge and Data Engineering* 10.6 (1998): 872-888.

[6] Allani, Olfa, et al. "A knowledge-based image retrieval system integrating semantic and visual features." *Procedia Computer Science* 96 (2016): 1428-1436.

[7] Qayyum, Zia Ul. "Image retrieval through qualitative representations over semantic features." *Multimedia Tools and Applications* 74.6 (2015): 1935-1959.

[8] Majeed, Saqib, Zia ul Qayyum, and Sohail Sarwar. "SIREA: Image retrieval using ontology of qualitative semantic image descriptions." *2013 International Conference on Information Science and Applications (ICISA)*. IEEE, 2013.

[9] Özaydın, Umut, Theodoros Georgiou, and Michael Lew. "A comparison of cnn and classic features for image retrieval." *2019 International Conference on Content-Based Multimedia Indexing (CBMI)*. IEEE, 2019.

[10] Mezaris, Vasileios, Ioannis Kompatsiaris, and Michael G. Strintzis. "Region-based image retrieval using an object ontology and relevance feedback." *EURASIP Journal on Advances in Signal Processing* 2004.6 (2004): 1-16.

[11] Jun, HeeJae, et al. "Combination of multiple global descriptors for image retrieval." *arXiv preprint arXiv:1903.10663* (2019).

[12] Yasmin, Mussarat, et al. "An Efficient Content Based Image Retrieval using El Classification and Color Features." *Journal of applied research and technology* 12.5 (2014): 877-885.

[13] Gordo, Albert, et al. "End-to-end learning of deep visual representations for image retrieval." *International Journal of Computer Vision* 124.2 (2017): 237-254.

[14] Qayyum, Zia Ul. "Image retrieval through qualitative representations over semantic features." *Multimedia Tools and Applications* 74.6 (2015): 1935-1959.

[15] Ladhake, Siddarth. "Semantic image analysis for intelligent image retrieval." *Procedia Computer Science* 48 (2015): 192-197.

[16] Philbin, James, et al. "Object retrieval with large vocabularies and fast spatial matching." *2007 IEEE conference on computer vision and pattern recognition*. IEEE, 2007.

[17] Zhang, Yimeng, Zhaoyin Jia, and Tsuhan Chen. "Image retrieval with geometry-preserving visual phrases." *CVPR 2011*. IEEE, 2011.

[18] Lindeberg, Tony. "Scale invariant feature transform." (2012): 10491.

[19] Zhou, Bolei, et al. "Learning deep features for discriminative localization." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

[20] Eitel, Andreas, et al. "Multimodal deep learning for robust RGB-D object recognition." *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2015.

[21] Shekhar, Ravi, and C. V. Jawahar. "Word image retrieval using bag of visual words." *2012 10th IAPR International Workshop on Document Analysis Systems*. IEEE, 2012.

[22] Tajeripour, Farshad, Mohammad Saberi, and Shervan Fekri Ershad. "Developing a novel approach for content based image retrieval using modified local binary patterns and morphological transform." *Int. Arab J. Inf. Technol.* 12.6 (2015): 574-581.