CBIR System for Remote Sensing Images

Content-Based Image Retrieval (CBIR) systems have revolutionized the way images are organized and retrieved in large databases, particularly in the domain of remote sensing. CBIR systems facilitate the retrieval of images not through traditional metadata or annotations but by analyzing the content of the images themselves. This approach is particularly useful in remote sensing where the volume of data is massive and the need for precise and efficient retrieval methods is critical.

Remote sensing technologies capture detailed imagery of the Earth's surface, which is utilized across various applications, from agricultural monitoring and urban planning to environmental conservation. The challenge lies in effectively managing and retrieving relevant images from these extensive datasets. Traditional methods that rely on manual tagging and metadata are often impractical due to the sheer volume and complexity of the data involved.

The reviwed papers:

1. Visual search over billions of aerial and satellite images(Keisler, 2019) :

Methodologies

They adapted a 50-layer ResNet CNN, initially trained on ImageNet, for their specific tasks. This network processes the aerial and satellite images to extract meaningful feature vectors that can be used for visual search.
Binary Features: To optimize data storage and computation, the feature vectors generated by the CNN were converted into binary formats. This involved modifying the network to produce almost-binary features, which were then thresholded to binary values. This adaptation was crucial for reducing the data footprint and facilitating fast retrieval.
Customization for Aerial and Satellite Imagery: For aerial imagery, like that from the National Agriculture Imagery Program (NAIP), the network was fine-tuned using images classified into 130 different object classes based on OpenStreetMap data. For the coarser resolution Landsat 8 data, an unsupervised autoencoder approach was used to compress features.

Search Methodologies:
Direct Search: This brute-force method involves computing Hamming distances between the query image's feature vector and all others in the dataset, identifying the closest matches. This method, while straightforward, is computationally expensive and was primarily used with the smaller Landsat 8 dataset.
Hash-based Search: For the larger NAIP dataset, a hash-based method was employed. This method used bit sampling for locality-sensitive hashing, where multiple hash tables store subsets of the binary feature vectors, allowing for quicker retrieval of approximate nearest neighbors.

Datasets
Aerial Imagery:
National Agriculture Imagery Program (NAIP): Provides aerial imagery across the continental United States, serving as one of the primary datasets for the visual search system. The imagery from NAIP was used to train the CNN with supervised learning, classifying images into various land-use categories.
Satellite Imagery:
Landsat 8 data, used to implement and test the visual search system on a global scale. This dataset was handled with an unsupervised learning approach due to its coarse resolution, which does not support detailed object classification.

The visual search system developed by Keisler et al. integrates these methodologies and datasets to provide a tool capable of searching billions of images in real-time. The binary encoding of image features and the efficient search algorithms allow users to quickly find images similar to a given query image, covering applications from rapid exploration of geographic locations to aiding in dataset creation for further machine learning tasks.

2. AiTLAS: Artificial Intelligence Toolbox for Earth Observation(Dimitrovski, 2023):

Methodologies

Deep Learning Models: This benchmark suite evaluates over 500 models derived from ten state-of-the-art architectures, focusing on multi-class and multi-label classification tasks in Earth Observation (EO).
Training Approaches:
The models are assessed both when trained from scratch and via transfer learning using pre-trained models, which is common in practical applications.
To promote reproducibility and allow further development by the research community, all experimental resources, including models, configurations, and dataset processing details, are made publicly available.

The benchmark includes 22 diverse datasets, which vary significantly in size, complexity, and the type of remote sensing data they contain, from high-resolution aerial images to multi-spectral satellite images.
These datasets are utilized for both multi-class and multi-label classification tasks, providing a comprehensive test bed for evaluating the versatility of deep learning models in handling various types of classification scenarios.

Detailed Dataset Characteristics:
All datasets used are open-access, which supports the transparency and applicability of the benchmark. The datasets are relevant to common EO tasks, enhancing the practical utility of the benchmark findings.
To ensure consistency in model training and evaluation, predefined splits are utilized where available, and new splits are created for datasets lacking them.

Benchmark Arena offers a robust platform for assessing the performance of state-of-the-art deep learning models in the field of Earth Observation. By providing a standardized and reproducible benchmarking environment, this work enables researchers and practitioners to evaluate and refine deep learning approaches effectively. The availability of this benchmark suite as an open-source resource ensures that it can serve as a foundational tool for ongoing and future research in remote sensing image classification

3. State of the art content based image retrieval techniques using deep learning: a survey(Kapoor, 2021)

Methodologies

This survey paper explores a variety of deep learning-based Content-Based Image Retrieval (CBIR) methods. It provides an extensive examination of the approaches used to bridge the semantic gap between the low-level features captured by machines and the high-level semantics anticipated by humans. Key methodologies discussed include:

Deep Learning Approaches for CBIR:

Binary Multi-scale Multi-pooling Approach: Generates deep binary codes based on the responses of feature maps from different deep convolutional layers. This method improves search accuracy by using a late fusion approach to combine the retrieval results from multiple layers.
Semantic Assisted Visual Hashing (SAVH): Integrates visual features with textual semantics using offline learning to enhance the retrieval performance by addressing the semantic gap more effectively.
Deep Hashing Based on Classification and Quantization Error (DHCQ): Focuses on minimizing both quantization and classification errors during the hash code generation process, aiming for high retrieval performance with efficient data compression.
Use of Advanced Pooling Techniques: Including Spatial Maximal Activator Pooling and Dynamic Late Fusion to enhance feature extraction and retrieval accuracy.

Evaluation Metrics and Similarity Measures:

The survey discusses various evaluation metrics used to assess CBIR systems, such as precision, recall, and mean average precision (mAP).
Highlights different similarity measures used in CBIR systems like Hamming distance, Euclidean distance, and cosine similarity, which are critical in defining how image similarity is quantified.

Datasets

INRIA Holidays: A dataset commonly used for evaluating image retrieval algorithms, consisting of personal holiday photos.
Oxford Buildings: Contains images of buildings from different landmarks designed to test image retrieval effectiveness.
UKBench: Includes images grouped into sets of four that are visually similar, used to test the retrieval rate and quality of CBIR systems.

NUS-WIDE: A dataset used for evaluating multi-label CBIR systems, containing images tagged with multiple labels from a wide range of categories.
MIRFLICKR-1M: A large dataset used for scalability testing in CBIR systems, offering a diverse range of images sourced from Flickr.

This survey integrates a wide range of deep learning techniques and evaluations to provide a comprehensive view of current capabilities and future directions in CBIR. By analyzing different methodologies and their applicability to various datasets, the paper aids in understanding the strengths and weaknesses of current approaches, guiding future research towards addressing the challenges in the field.

4. Image retrieval from remote sensing big data: A Survey(Li, 2020)

Methodologies

The paper categorizes the methods of RS image retrieval into several distinct types: conventional content-based RS image retrieval, hashing-based RS image retrieval, cross-modal RS image retrieval, and interactive RS image retrieval. Each category encompasses various algorithms and techniques tailored to the unique challenges of retrieving images from vast remote sensing databases.

Conventional Content-Based Image Retrieval:

Discusses various techniques for representing image features, from low-level spectral and texture features to high-level semantic features using deep learning.
Feature Indexing: Focuses on efficiently indexing these features to facilitate quick retrieval, utilizing tree-based, clustering-based, and hashing-based indexing methods.
Examines methods to measure similarities between images, essential for determining the relevance of search results to the query image.

Hashing-Based RS Image Retrieval:

Emphasizes the importance of converting high-dimensional image descriptors into compact binary codes to speed up the retrieval process and reduce storage requirements. The discussion includes both supervised and unsupervised hashing techniques.

Cross-Modal RS Image Retrieval:

Explores techniques for retrieving images based on inputs that may not be images themselves, such as text descriptions or sketches, highlighting the need for models that can interpret and bridge different data modalities.

Interactive RS Image Retrieval:

Incorporates user feedback to refine search results iteratively, enhancing the relevance of retrieved images by adjusting the retrieval process based on user interactions.

The survey provides a detailed discussion on various datasets that are commonly used in remote sensing image retrieval research, underscoring the diversity and complexity of data in the field.

Highlights include high-dimensional datasets with extensive label sets and varying resolutions, emphasizing the challenges in handling and processing such large-scale data effectively.

Evaluation Metrics:

Reviews common metrics used to assess the performance of RS image retrieval systems, such as precision, recall, and the mean average precision (mAP), which are crucial for evaluating the effectiveness of different retrieval methodologies.

This survey paper integrates a wide array of techniques and approaches under the umbrella of RS image retrieval, providing a comprehensive look at how these methods are applied in practice. By examining both the technological and application-oriented aspects of RS image retrieval, the paper helps pave the way for future innovations in the field. It also points out several research gaps and emerging trends that could further improve retrieval performance and system efficiency.

Conclusion:

The objective of our project is to analyze the impact of employing deep learning methods and artificial neural networks in designing a Content-Based Image Retrieval (CBIR) system. This system is tailored for handling multiscale images sourced from various origins with different spatial resolutions. Initially, I prepared and analyzed datasets specific to this task. Subsequently, I utilized pretrained models for feature extraction to visually and statistically evaluate performance. I also incorporated additional datasets to enrich our analysis. While employing hash coding to accelerate the search process for extracted features—similar to the systems discussed in the referenced papers—was considered, it was not the primary focus of our project.

Following this, we fine-tuned the pretrained models based on the datasets mentioned and documented the results. The codes and implementations are available in our GitHub repository.