Fundamentals of image Representation

Basics of Image Processing, Feature Extraction, Descriptors, Feature Vectors, Indexing, and Similarity Measures

Dr. Muhammad Sajjad

R.A: Kaleem Ullah

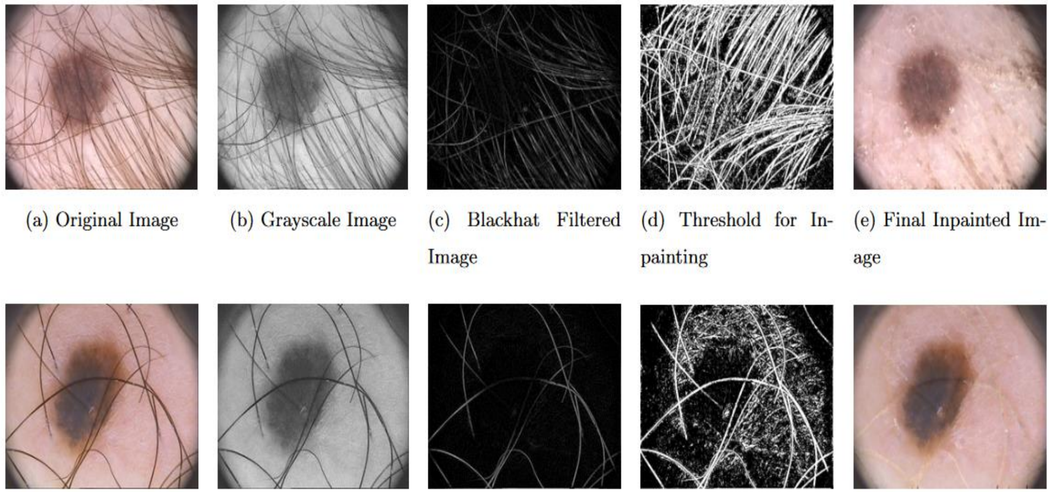
R.A: Muhammad Ayaz

Overview

- Basics of Image Processing
- Feature Extraction
- Image Descriptors and Feature vectors
- Indexing Techinques for images and videos
- Similarity Measures for images and videos
- Applications
- Challenges and Future Treds

Basics of Image Processing

- Image processing stands as a pivotal field, a dynamic domain where images undergo meticulous manipulation and analysis.
- Its essence lies in the extraction of valuable information or the enhancement of visual quality.
- This fundamental process is not confined to a specific realm; instead, it permeates across various domains, playing an integral role in the tapestry of numerous technological advancements.

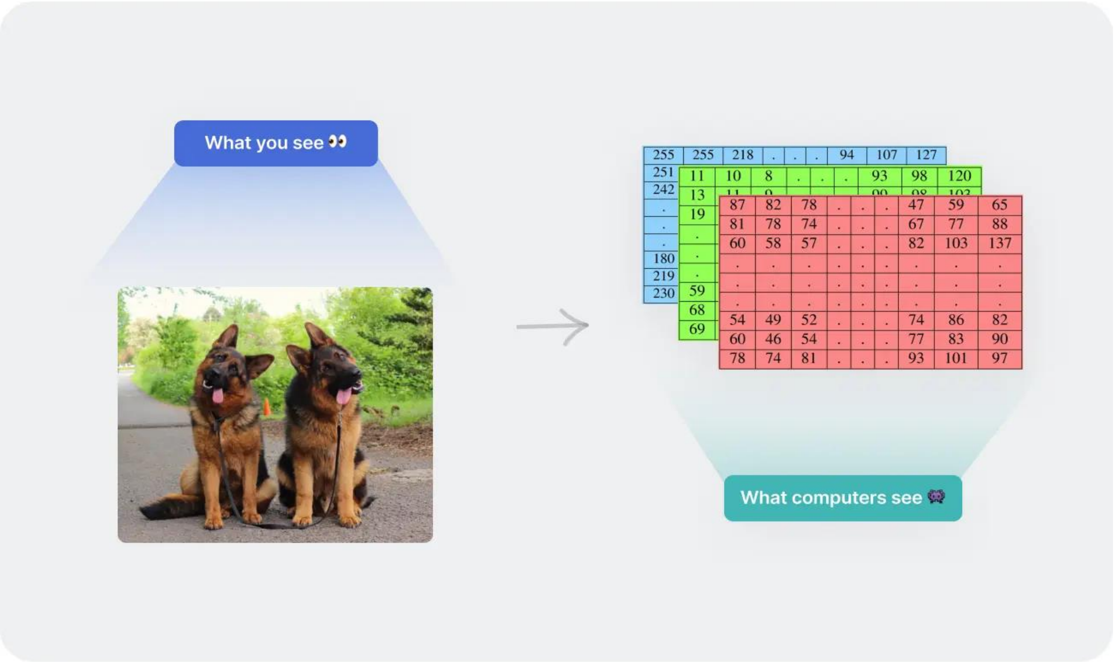


(a) Original Image (b) Grayscale Image (c) Blackhat Filtered Image (d) Threshold for In-painting (e) Final Inpainted Image

(f) Original Image (g) Grayscale Image (h) Blackhat Filtered Image (i) Threshold for In-painting (j) Final Inpainted Image

How Machines "See" the image:

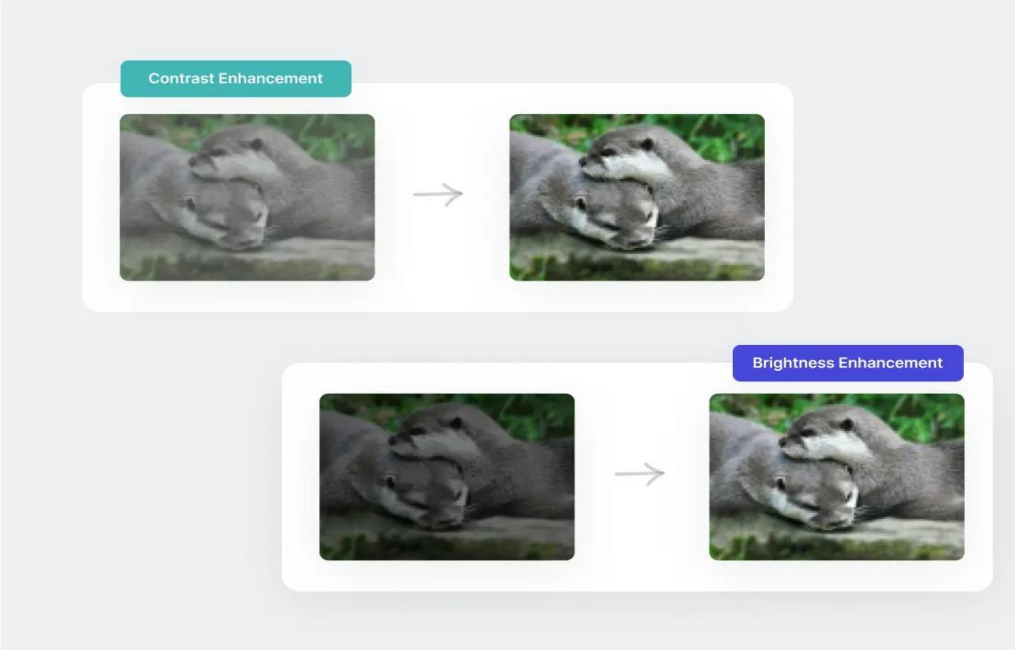
- Digital images are represented as 2D or 3D matrices, where each pixel value corresponds to the pixel's intensity.
- Typically, we work with 8-bit images, where the intensity ranges from 0 to 255.
- Understanding how machines perceive images is fundamental for effective image processing.

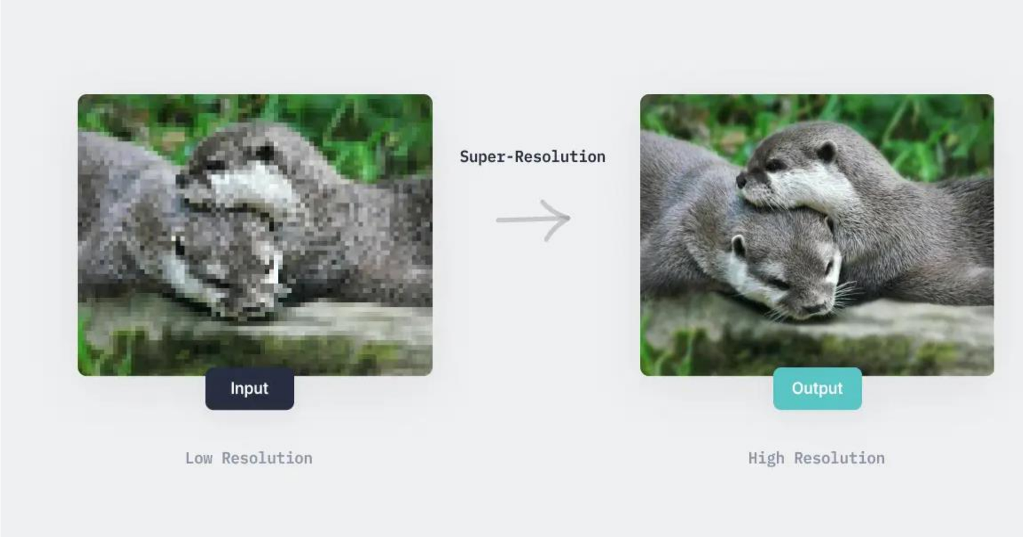


Phases of Image Processing

- Image Acquisition
  - The process begins with capturing an image using a camera.If the camera output is not already digitized, an **analog-to-digital converter** is used to convert it into a digital format.

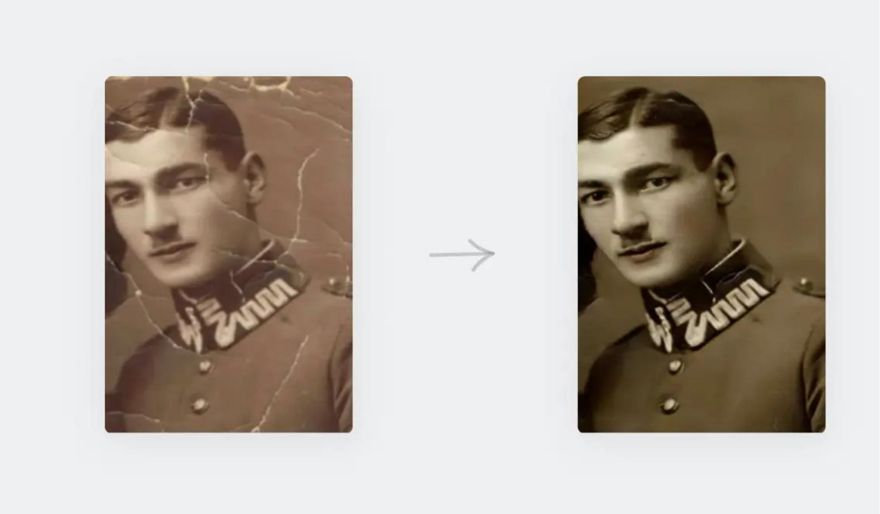
- Image Enhancement
  - In this step, the acquired image is manipulated to meet the requirements of the specific task for which the image will be used.
  - Such techniques are primarily aimed at highlighting the hidden or important details in an image, like contrast and brightness adjustment, etc. Image enhancement is highly subjective in nature.



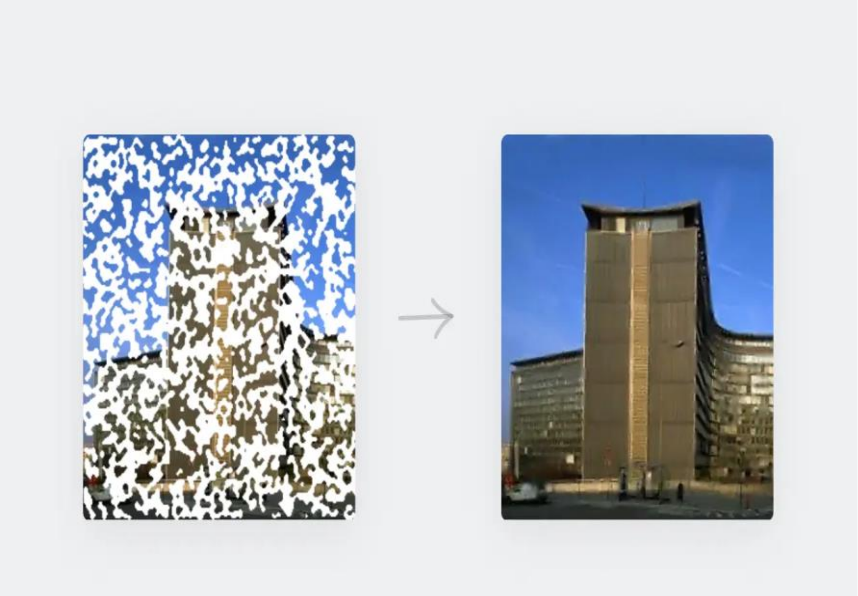


Phases of Image Processing

- **Image Restoration**
- The quality of images could degrade for several reasons, especially photos from the era when cloud storage was not so commonplace. For example, images scanned from hard copies taken with old instant cameras often acquire scratches on them.

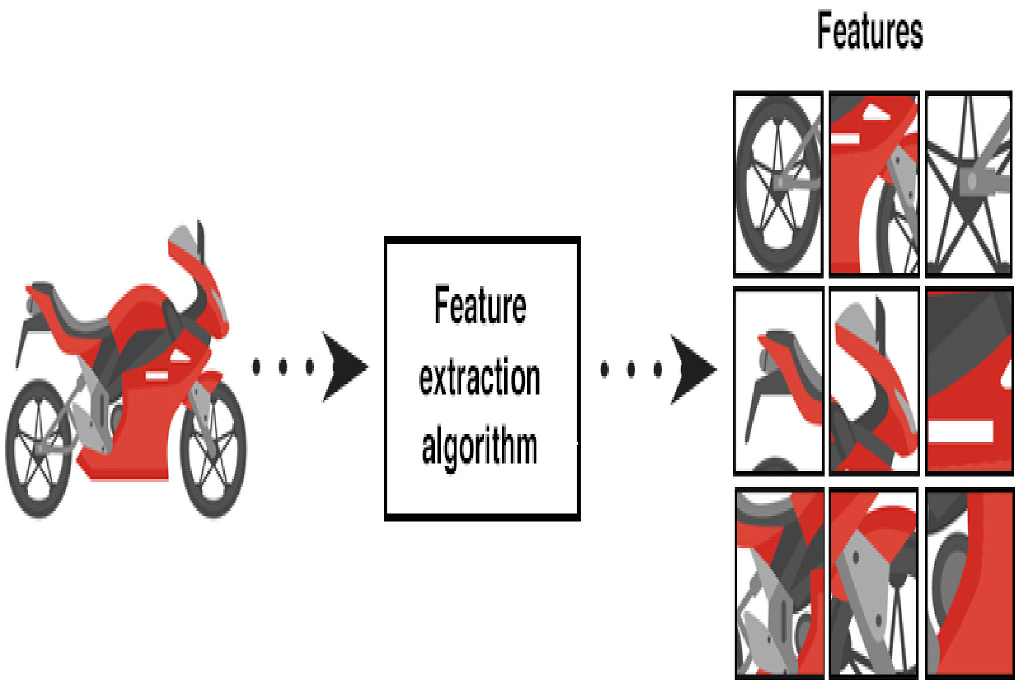


- **Image Segmentation**
  - This step involves partitioning an image into different key parts to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze.
  - Image segmentation allows for computers to put attention on the more important parts of the image, discarding the rest, which enables automated systems to have improved performance.



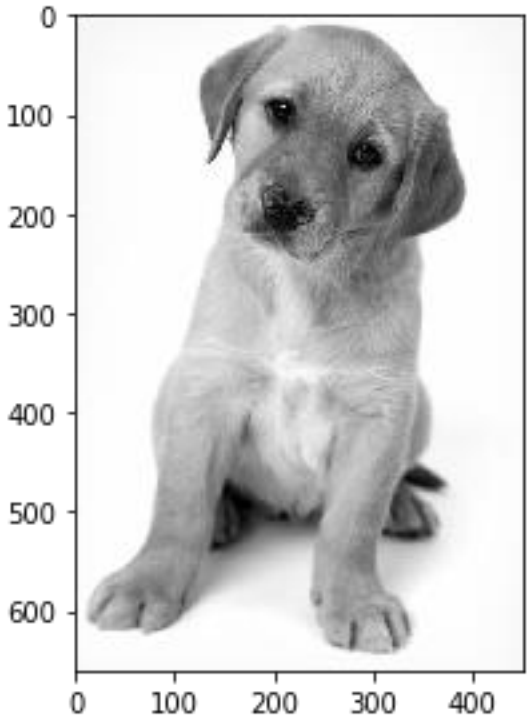
Feature Extraction

- Feature extraction is the process of selecting, transforming, and representing relevant information or patterns from raw data, often in the context of signal processing or pattern recognition.
- The primary goal is to reduce the dimensionality of the data while retaining essential information, making it more manageable for analysis and modeling.
- In the context of different applications, feature extraction involves selecting or transforming specific characteristics, patterns, or attributes of the data that are relevant to the task at hand.
- These features are then used as input for machine learning models. Feature extraction is crucial in scenarios where the raw data may be too complex, noisy, or redundant, making it challenging for algorithms to effectively learn and generalize.



Techniques to Extract Features from Image

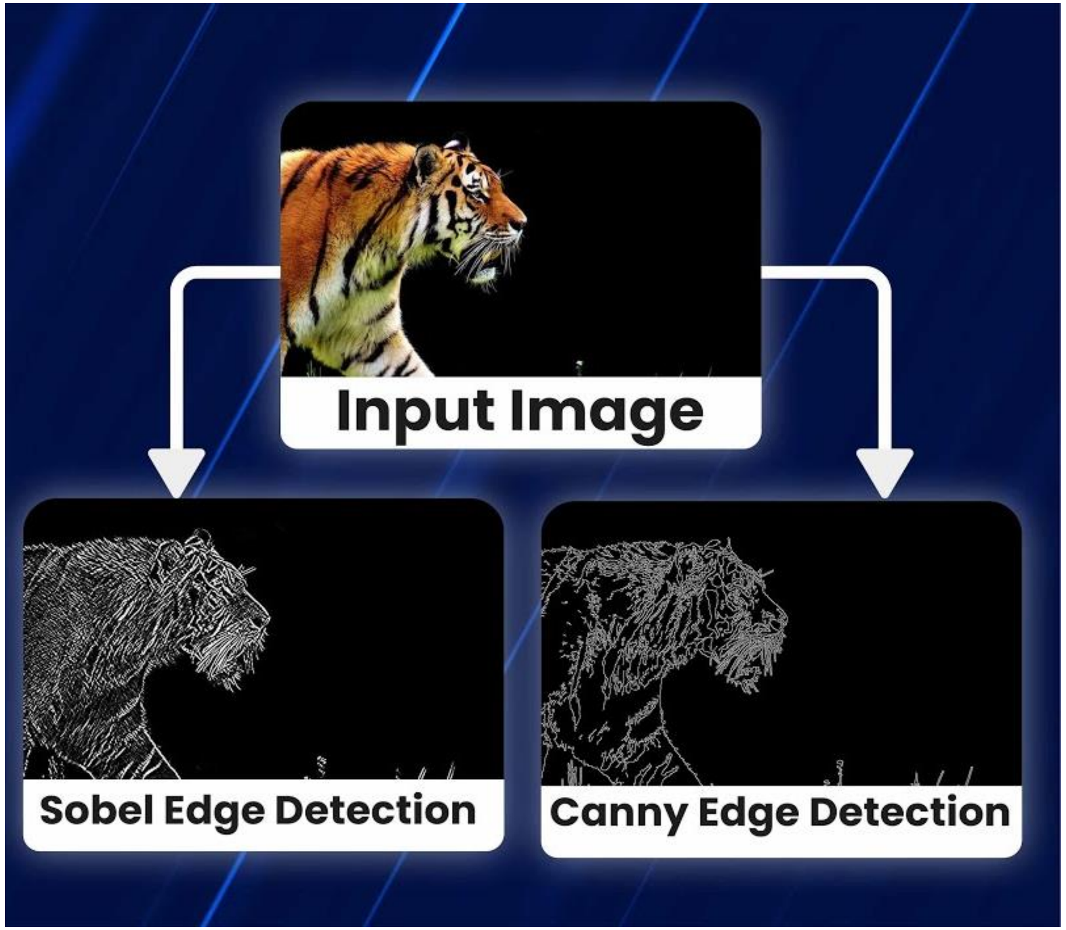
- **Grayscale Pixel Values as Features**
- In this method, we convert the image to grayscale, resulting in a single channel image.
- Each pixel's intensity value becomes a feature. For an image of dimensions **n x m**, we have **n x m** features.
- These grayscale pixel values can be used as input features for machine learning models.
- For example, consider an image of the number 8. When we examine it closely, we notice that it is composed of small square boxes (pixels).
- Machines store this image as a matrix of pixel values, where the dimensions represent the number of pixels (height x width). These pixel values serve as our features.



(297000,)
array([0.96470588, 0.96470588, 0.96470588, ..., 0.96862745, 0.96470588, 0.96470588])

Techniques to Extract Features from Image

- **Extracting Edge Features**:
- Edge detection techniques identify abrupt changes in intensity within an image.
- These changes often correspond to object boundaries or edges.
- Common edge detection filters include the Sobel, canny, and Roberts operators. By extracting edges, we can capture important structural information from the image
- The Sobel operator computes an **approximation of the gradient** of the image intensity function. At each point in the image, it provides either the corresponding gradient vector or the **norm** of this vector



Techniques to Extract Features from Image

- **Histogram of Oriented Gradients (HOG)**:
- HOG is a powerful feature extraction method used for object detection and recognition.
- It computes histograms of gradient orientations within local image regions.
- These histograms capture texture and shape information.
- HOG features are particularly effective for detecting objects with varying shapes and appearances

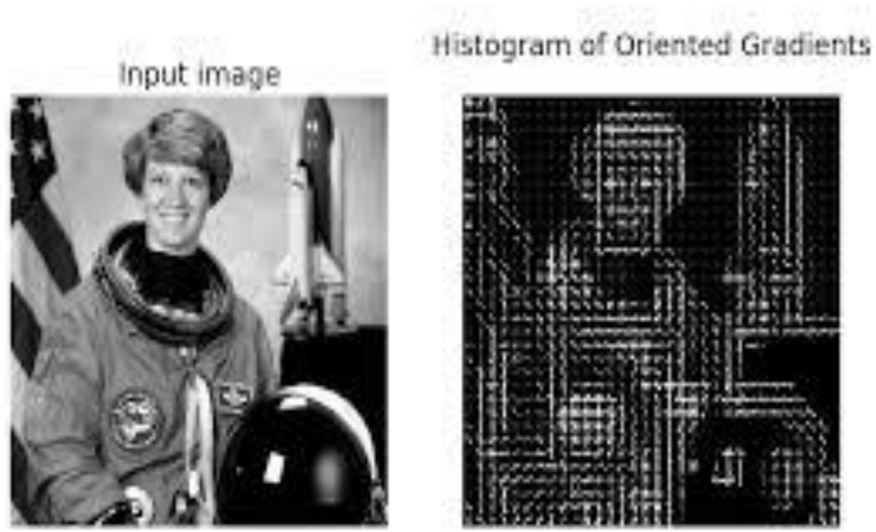
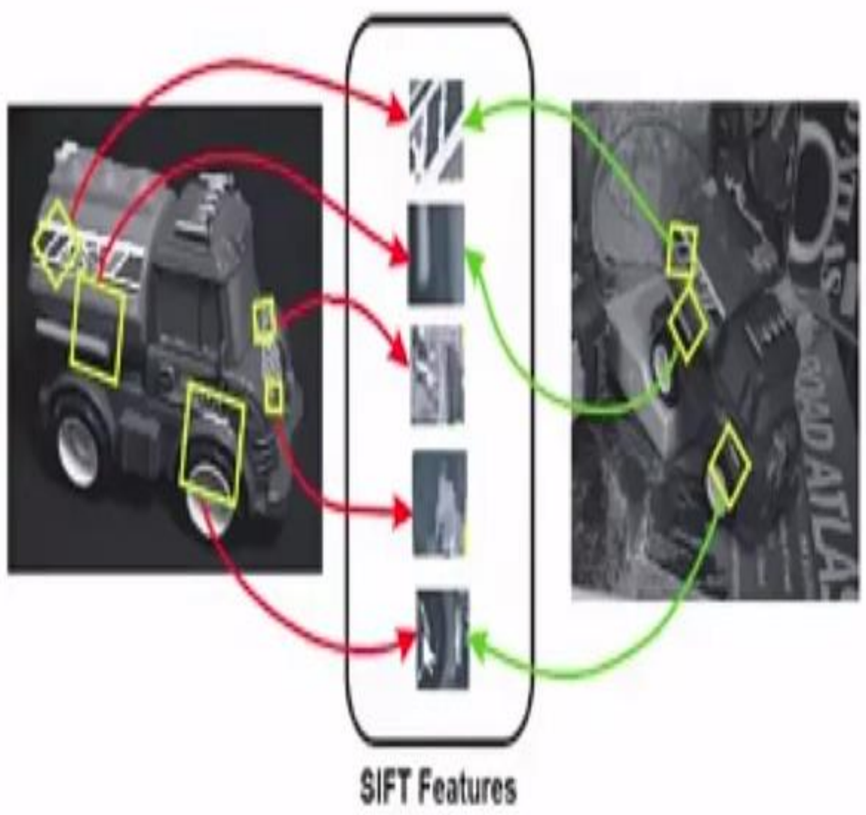


Image Descriptors and feature vectors

- **Image descriptors** and **Feature vectors** play a crucial role in computer vision, image processing, and machine learning. They represent essential information about the content of an image and are used for tasks such as image recognition, object detection, and image retrieval.
- **Image descriptors**
  - Image descriptors are representations of visual information extracted from an image. They capture various characteristics, patterns, or structures present in the image.
  - **Purpose:** Descriptors aim to provide a compact and meaningful representation of the image content, allowing for efficient analysis and comparison between images
  - **Examples:** Descriptors can be local or global.
  - Local descriptors focus on specific regions or keypoints in an image and include SIFT (Scale-Invariant Feature Transform), SURF (Speeded-Up Robust Features), and ORB (Oriented FAST and Rotated BRIEF).
  - Global descriptors represent the entire image, such as color histograms or texture descriptors

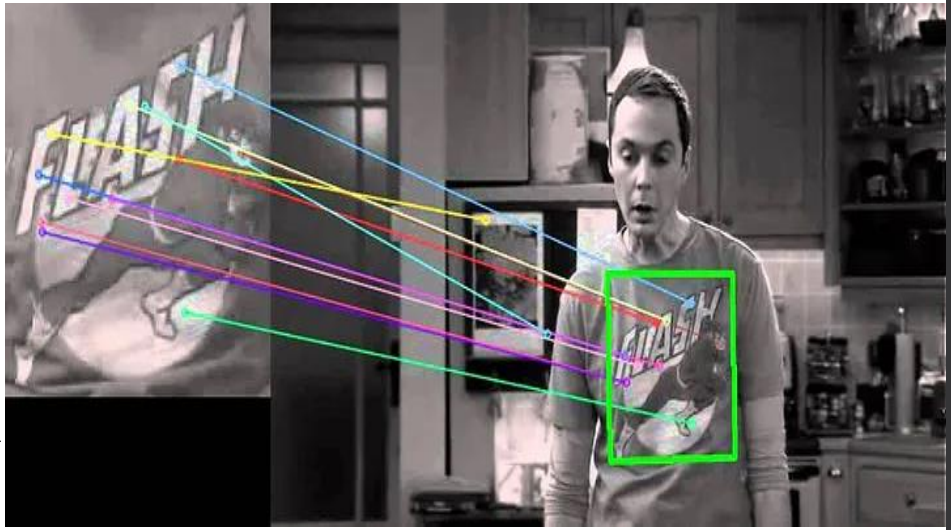
SIFT (Scale-Invariant Feature Transform):

- **Detection:** SIFT identifies distinctive keypoints in an image, which are regions that are invariant to scale changes and rotations.
- **Description:** Once keypoints are detected, SIFT generates descriptors that capture information about the local gradient orientations and magnitudes around each keypoint.
- **Invariance:** SIFT is designed to be robust to changes in scale, rotation, and illumination, making it suitable for various computer vision tasks such as image matching and object recognition.
- **Applications:** SIFT has been widely used in applications like image stitching, object recognition, and 3D reconstruction



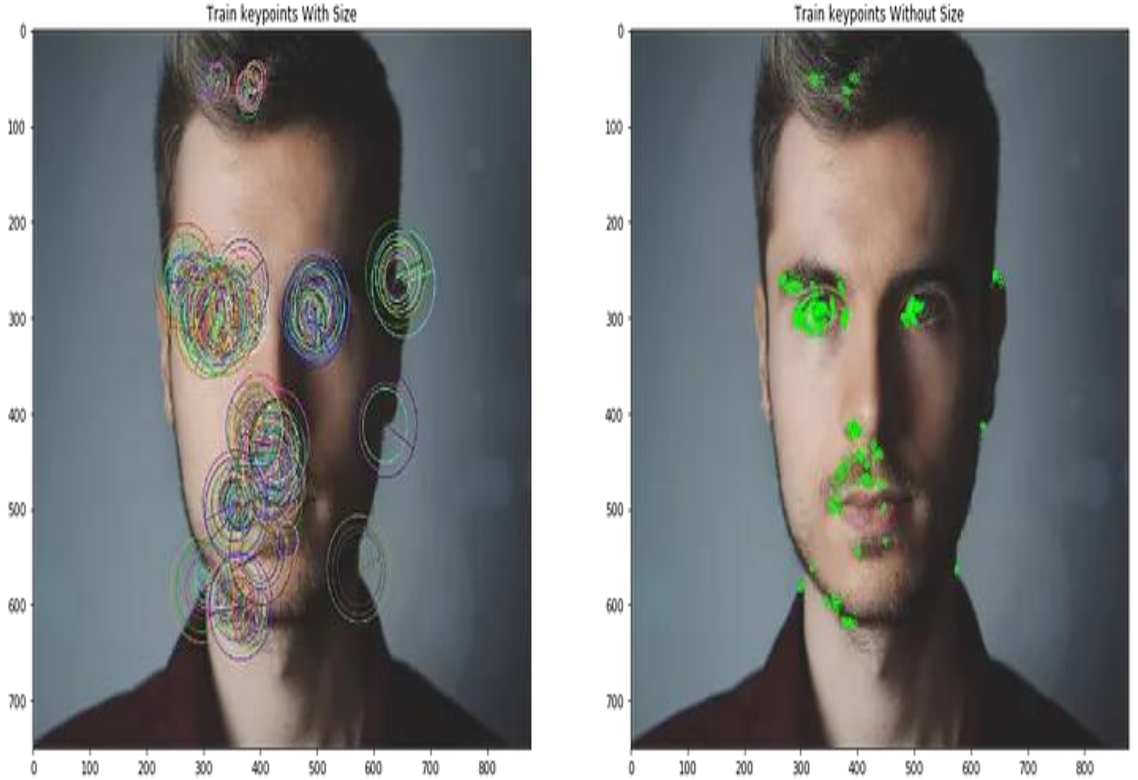
SURF (Speeded-Up Robust Features)

- The SURF method (Speeded Up Robust Features) is a fast and robust algorithm for local, similarity invariant representation and comparison of images.
- The main interest of the SURF approach lies in its fast computation of operators using box filters, thus enabling real-time applications such as tracking and object recognition
- **Efficiency:** SURF is computationally more efficient than SIFT, making it suitable for real-time applications.
- **Interest Points:** SURF uses a faster interest point detector and approximates the calculation of gradient information, contributing to its speed.
- **Applications:** SURF is commonly used in real-time object recognition, image stitching, and other applications where computational efficiency is crucial.



ORB (Oriented FAST and Rotated BRIEF)

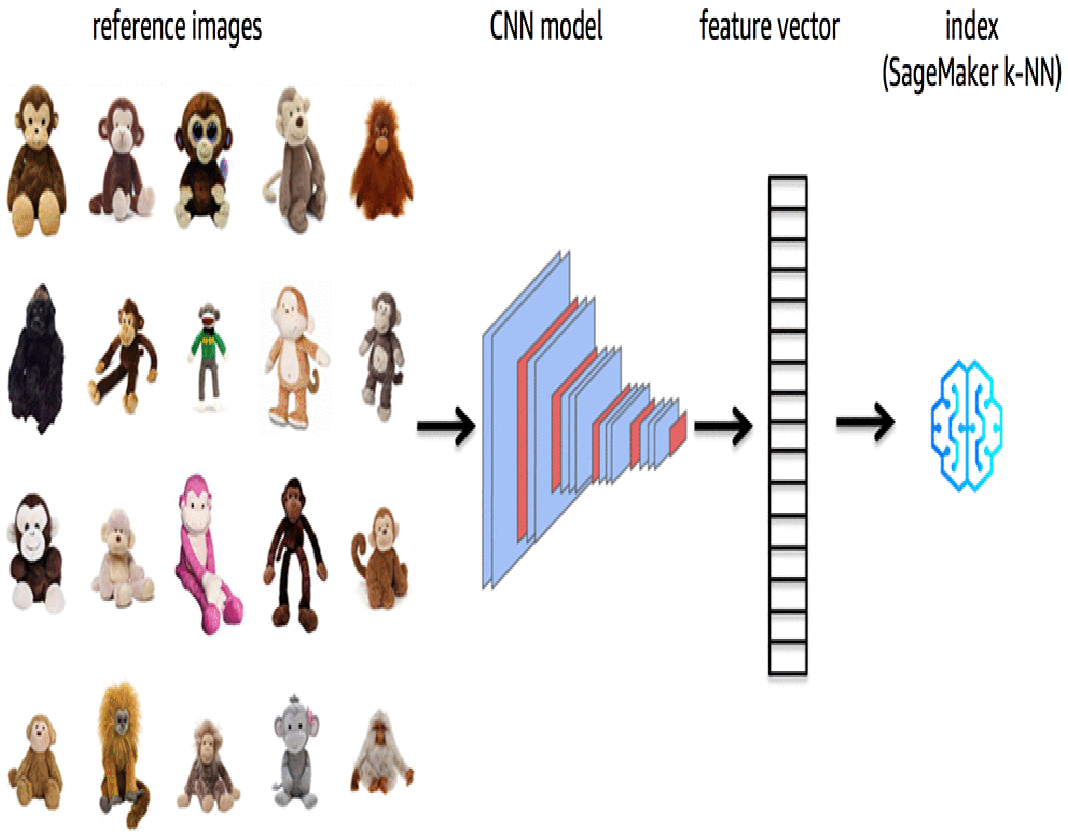
- ORB is basically a fusion of FAST keypoint detector and BRIEF descriptor with many modifications to enhance the performance.
- First it use FAST to find keypoints, then apply Harris corner measure to find top N points among them. It also use pyramid to produce multiscale-features.
- **Applications:**
- ORB is commonly used in applications where both speed and accuracy are crucial, such as in robotics, augmented reality, and real-time object recognition.
- It is often employed in scenarios where the computational demands of more complex feature descriptors, like SIFT and SURF, may be prohibitive.



Feature Vectors

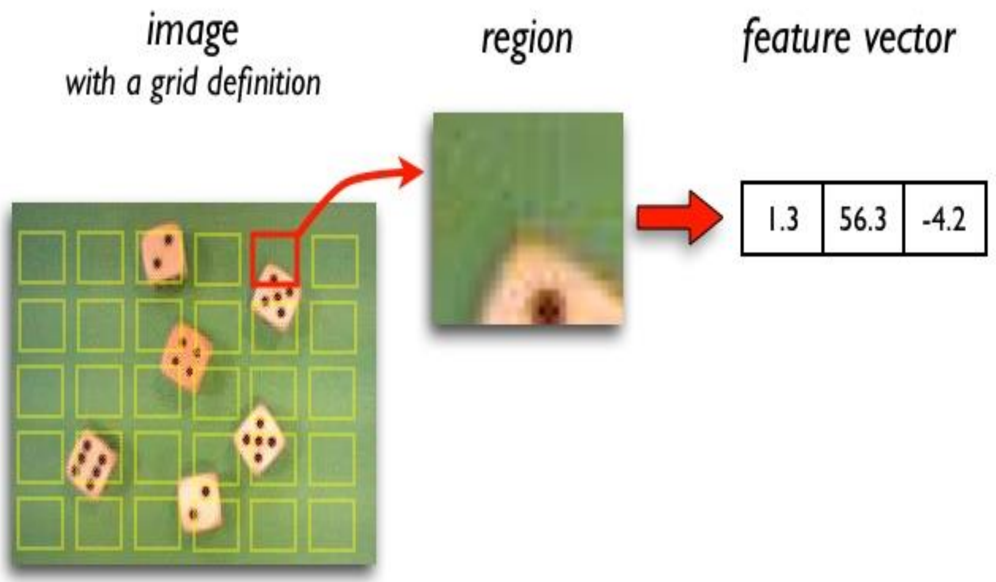
A feature vector is a numerical representation of an image derived from its descriptors. It is a one-dimensional array of values where each value corresponds to a specific feature or descriptor.

- **Purpose:** Feature vectors serve as input to machine learning algorithms, allowing the algorithms to learn patterns and make predictions based on the characteristics present in the images.
- **Examples:** In the context of image descriptors, a feature vector could be created by concatenating the values of various descriptors extracted from an image.
- For instance, a feature vector might include color histogram values, texture features, and local keypoints' descriptors.



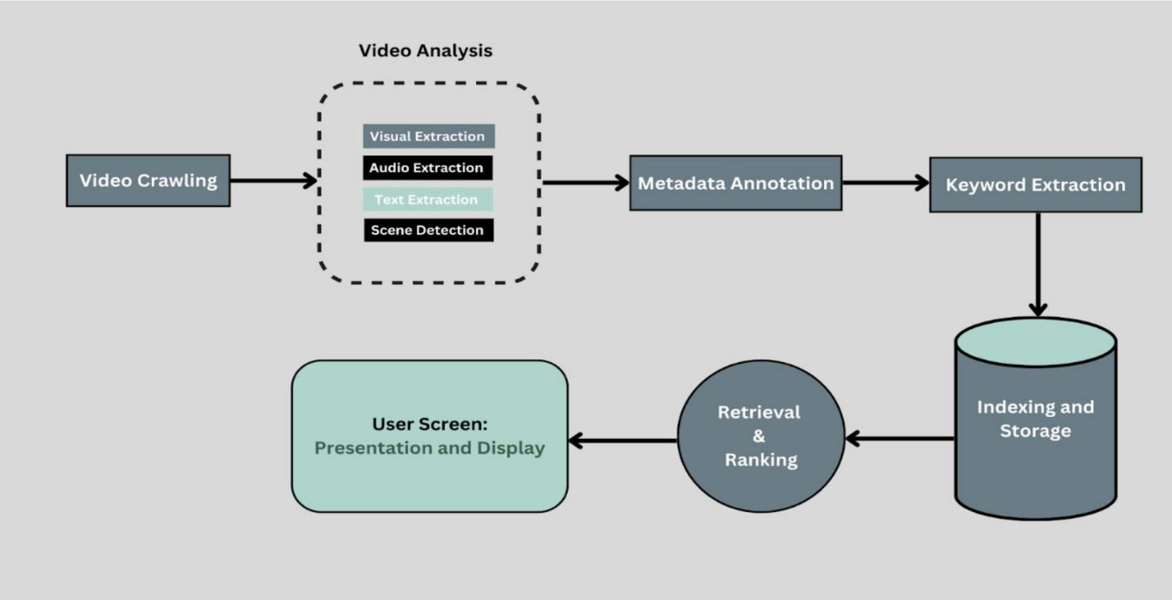
Feature Vectors

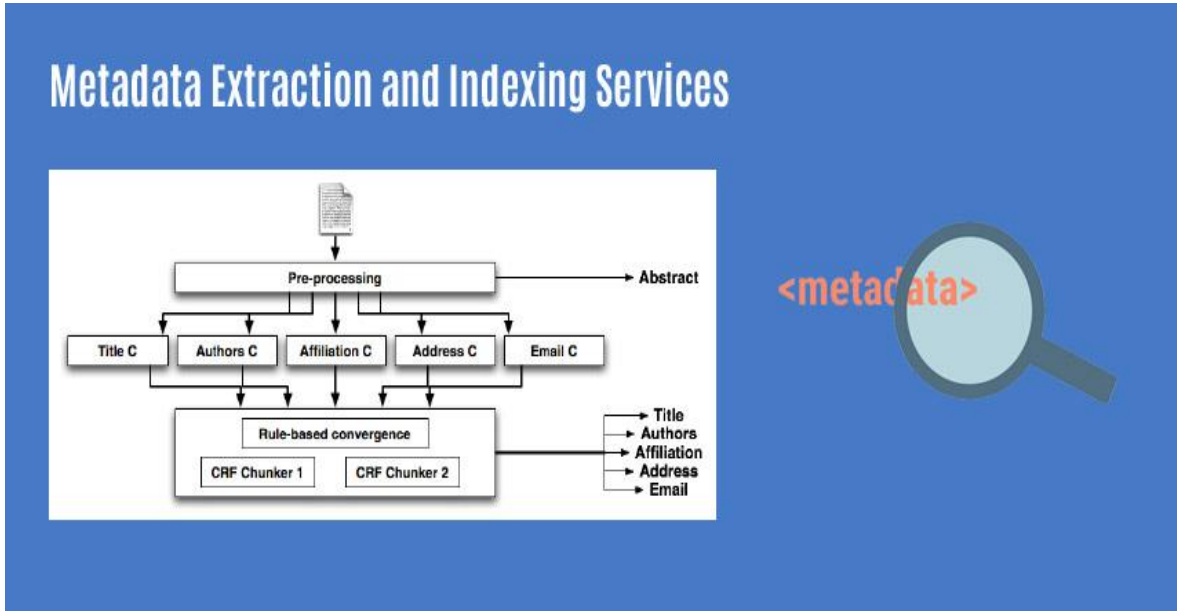
- **Process of Creating Feature Vectors:**
- **Extraction:** Use image descriptors (e.g., SIFT, HOG, color histograms) to extract relevant information from t image.
- **Normalization:** Ensure that the extracted features ar scaled or normalized to make them comparable and reduce the impact of varying scales.
- **Combination:** Combine the extracted descriptors into single feature vector. This can involve concatenating th values or using other methods based on the chosen descriptors.
- **Utilization:** The resulting feature vector can be used as input for machine learning models, facilitating tasks like classification, clustering, or retrieval.



Indexing Techinques for images and videos

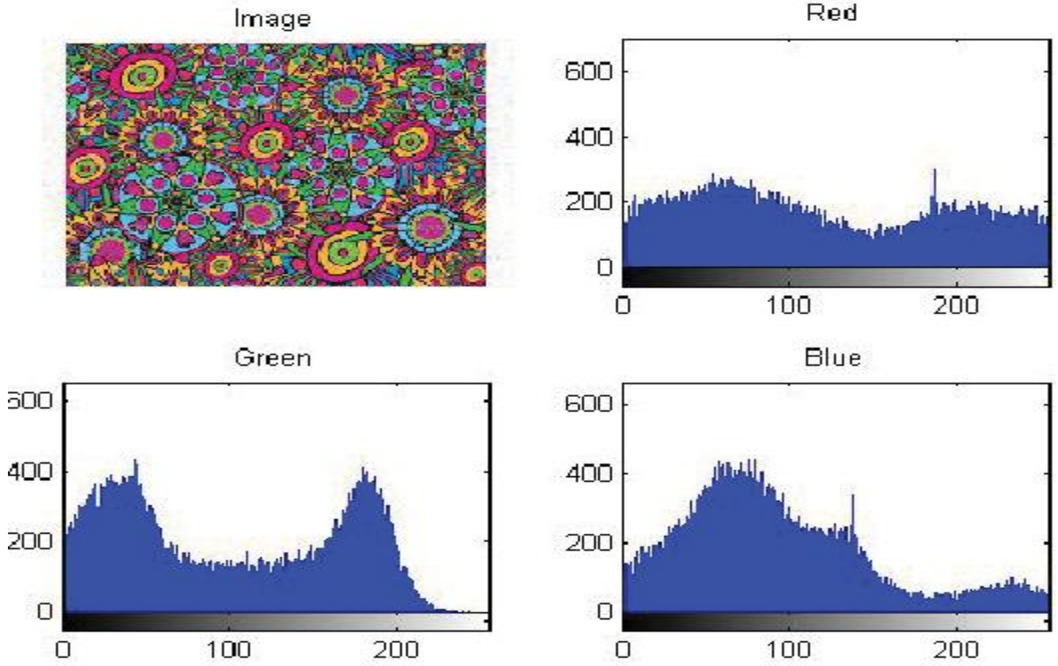
- Indexing techniques for images and videos involve methods for organizing, storing, and retrieving visual data efficiently.
- These techniques are crucial for managing large collections of multimedia data and facilitating quick access to specific content
- **Metadata Indexing:**
- Associate metadata with images or videos, such as keywords, tags, timestamps, and location information.
- **Advantages:** Enables fast retrieval based on descriptive information. Can be manually or automatically assigned.
- **Applications:** Commonly used in content management systems and social media platforms.





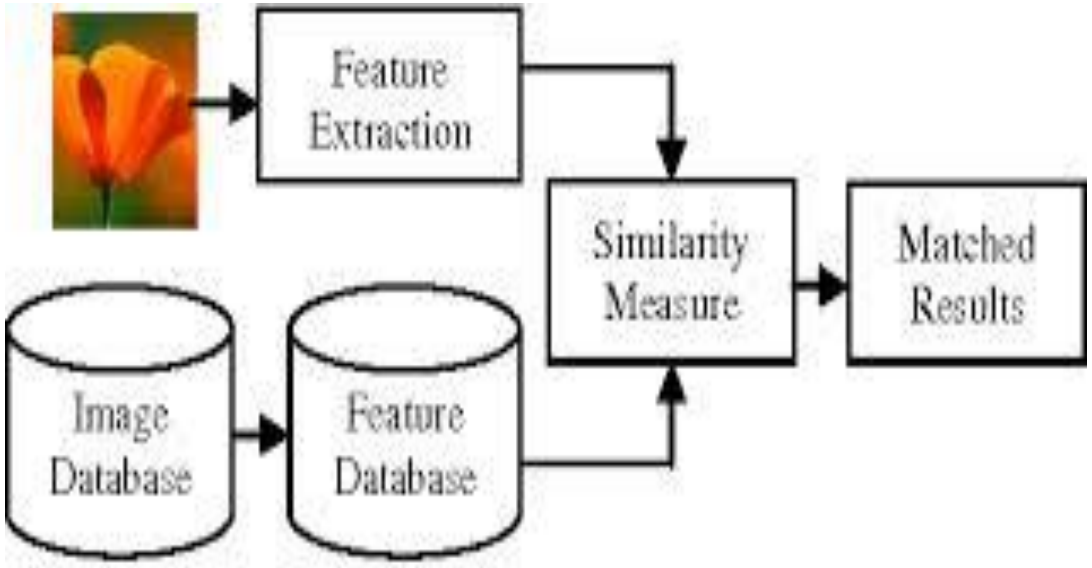
Indexing Techinques for images and videos

- **Color-Based Indexing:**
- Index images or video frames based on their color histograms or dominant color features.
- **Advantages:** Allows for quick retrieval of visually similar content based on color information.
- **Applications:** Useful for image search engines, where users seek images with specific color characteristics.



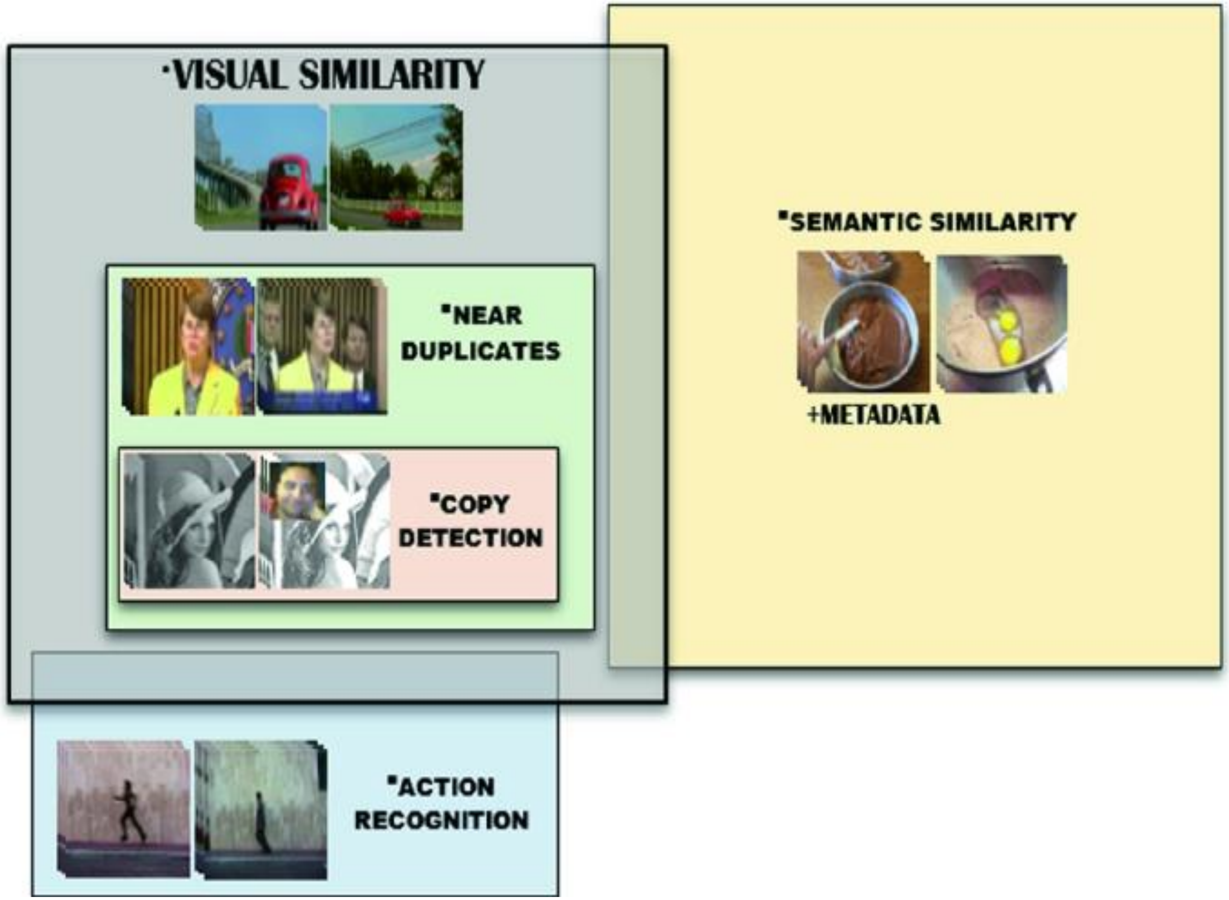
Indexing Techinques for images and videos

- **Feature-Based Indexing:**
- **Description:** Extract local features from images using techniques like SIFT, SURF, or ORB, and use these features for indexing.
- **Advantages:** Enables robust and distinctive feature matching, useful for object recognition and image retrieval.
- **Applications:** Widely used in computer vision applications for content-based image retrieval.



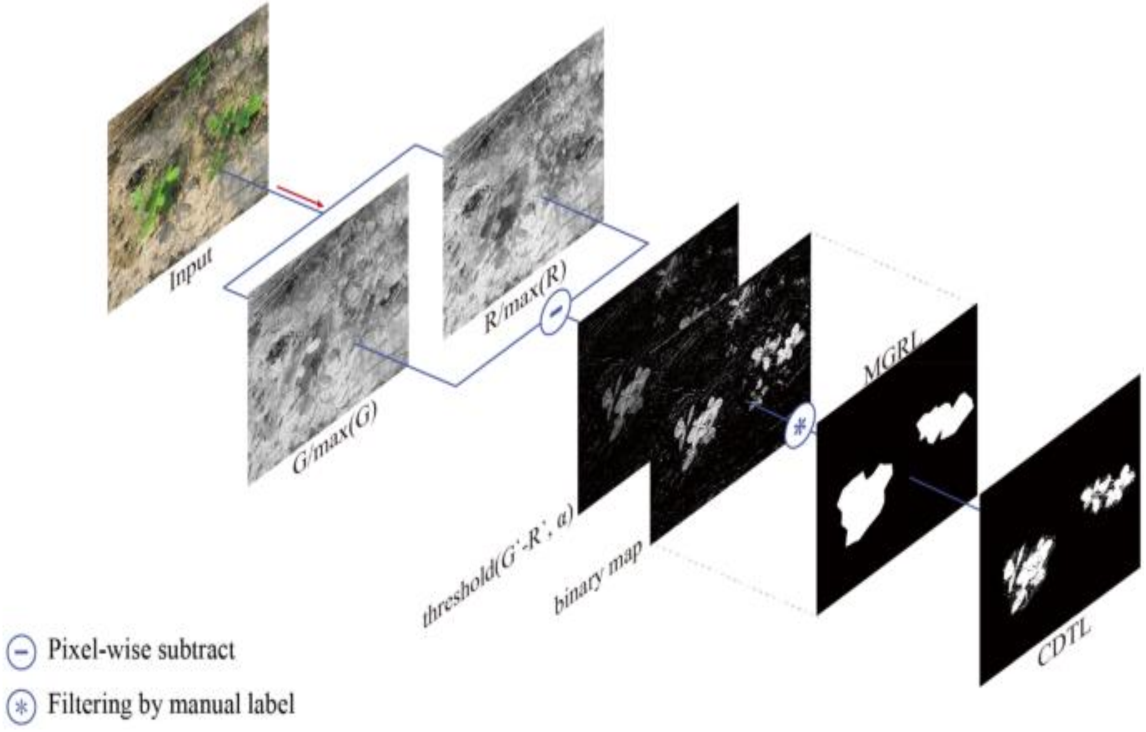
Similarity Measures for images and videos

- Similarity measures for images and videos are essential in various computer vision applications, such as content-based image retrieval, video summarization, and object recognition.
- Several techniques exist to quantify the similarity between images and videos.



Similarity Measures for images and videos

- **Image Similarity Measures**
- **Pixel-wise Distance**
- **Euclidean Distance**: Measures the straight-line distance between pixel values in two images.Manhattan **Distance**: Calculates the sum of absolute differences between pixel values.
- **Cosine Similarity**: Computes the cosine of the angle between pixel vectors

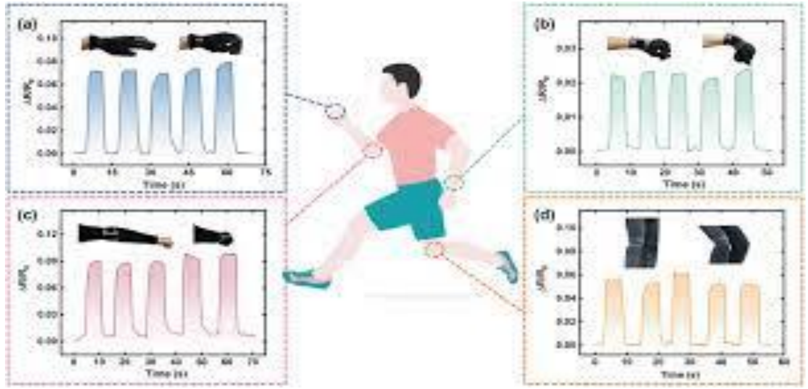


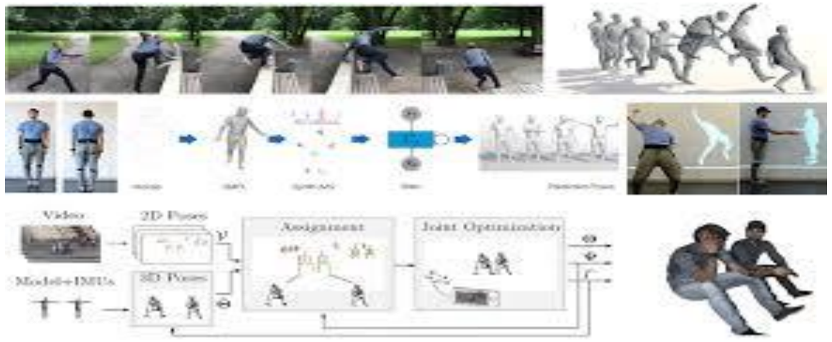
Similarity Measures for images and videos

- **Structural Similarity Index (SSI or SSIM):**
- Evaluates the structural information of images, considering luminance, contrast, and structure.
- SSIM values range from -1 to 1, with 1 indicating identical images.
- **Histogram-Based Measures:**
- Use color histograms to represent images and measure the similarity between their distributions.
- Methods include Bhattacharyya distance and Earth Mover's Distance (EMD).

Similarity Measures for images and videos

- **Video Similarity Measures**
- **Temporal Correlation:**
  - **Frame-wise Cross-Correlation**: Measures the similarity between corresponding frames in two videos.
  - **Dynamic Time Warping (DTW):** Aligns and measures the similarity between temporal sequences.
- **Motion-based Measures:**
  - **Optical Flow-based Measures**: Compares the motion patterns between frames.
  - **Trajectory Matching:** Compares the paths followed by objects or features over time.





Applications

- **Computer Vision:**
- **Object Recognition:** Image representations are used for training models to recognize and classify objects within images. This is essential in applications like autonomous vehicles, security systems, and robotics.
- **Image Segmentation:** Representing images helps in segmenting them into different regions or objects, allowing for more granular analysis and understanding.

Applications

- **Medical Imaging:**
  - **Diagnostic Imaging:** Image representation is fundamental in medical diagnostics, aiding in the detection of abnormalities in X-rays, MRIs, CT scans, and other imaging modalities.
  - **Disease Classification:** Machine learning models use image representations to classify medical images and assist in diagnosing conditions.
- **Augmented Reality (AR) and Virtual Reality (VR):**
  - Image representations are used in AR and VR applications to create realistic and immersive virtual environments. This includes overlaying digital information on the real world (AR) or creating entirely virtual worlds (VR).

Challenges and Future Treds

Challenges:

- **Data Privacy and Ethical Concerns:**
  - As image representation and computer vision technologies advance, concerns about privacy and ethical use of visual data become more prominent. Striking a balance between innovation and protecting individuals' privacy is a challenge.
- **Bias and Fairness:**
  - Biases present in training data can lead to biased outcomes in image analysis, affecting certain demographics more than others. Addressing and mitigating bias is crucial for fair and equitable applications.
- **Interpretable AI:**
  - Making image representation models more interpretable is a challenge. Understanding how these models arrive at specific decisions is essential, especially in critical applications like healthcare and autonomous vehicles.

Challenges:

- **Robustness to Adversarial Attacks:**
  - Image representation models are susceptible to adversarial attacks, where small, carefully crafted changes to an image can lead to incorrect predictions. Developing robust models that can withstand such attacks is an ongoing challenge.
- **Computational Resources:**
  - Training and deploying sophisticated image representation models often require significant computational resources. Optimizing models for efficiency without compromising performance is a challenge, especially for real-time applications.

Future Trends:

- **Explainable AI:**
  - Continued focus on developing models that provide clear explanations for their decisions. Explainable AI is crucial for gaining trust and understanding in applications like healthcare and finance.
- **Multimodal Learning:**
  - Integration of multiple modalities, such as text, audio, and video, to enhance the understanding of complex scenarios. This trend allows AI systems to leverage diverse sources of information for more comprehensive analysis.
- **Continual Learning:**
  - Building models that can adapt and learn continuously from new data without forgetting previously acquired knowledge. Continual learning is essential for applications with evolving environments.

Future Trends:

- **Generative Models:**
  - Advancements in generative models like GANs (Generative Adversarial Networks) for creating and enhancing images. These models have applications in content creation, art, and virtual environments.
- **Edge Computing for Image Processing:**
  - Moving image processing tasks closer to the source by leveraging edge computing. This trend aims to reduce latency, increase efficiency, and address privacy concerns by processing data locally.
- **AI in Creativity:**
  - Expanding the role of AI in creative fields such as art, design, and content creation. AI-driven tools are likely to become more prevalent in assisting and augmenting human creativity.

Thank You