Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 6, 2025

CONNECTED REGIONS FORMATION FOR IMAGE CLASSIFICATION

Jasra Asma[1], Saroosh Jaffar[2], Sadia Latif [*3], Rana Muhammad Nadeem[4], Adnan Altaf[5]

[1]*Department of Information Technology. Bahauddin Zakaria University, Multan, Pakistan*
[2]*Institute of Computer Science and Information Technology, The Women University,Multan*
[*3]*Department of Computer Science, Bahauddin Zakaria University, Multan, Pakistan*
[4] *Department of Computer Science, Govt. Graduate College Burewala, Pakistan*
[5]*School of Computer Science and Communication Engineering, Jiangsu University, Zhenjiang, Jiangsu 212013, China*

[*3]sadialatifbzu@gmail.com

DOI: https://doi.org/10.5281/zenodo.15574429

Keywords
Key point detection, Image descriptors, Pixel intensity, Symmetric sampling, PCA, Feature extraction, Visual words, Precision, Recall, Grayscale image, L2 normalization, Isotropic and anisotropic filtering, Cascade matching. Machine learning

Article History
Received on 25 April 2025
Accepted on 25 May 2025
Published on 02 June 2025

Corresponding Author: * **Sadia Latif**

Abstract

Many computer vision applications rely on matching key points between images. Over recent decades, advancements in key-point detection algorithms have significantly improved both robustness and speed. However, there is an ongoing need for more compact descriptors and faster methods with higher classification accuracy. This work addresses this need by introducing a novel algorithm that formulates both a key-point detector and descriptor based on prominent image features.

The proposed detector focuses on identifying intensity-based corners and edges within grayscale images. This process involves detecting connected regions by analyzing pixel intensity ranges. Once bright and dark regions are identified, pixel intensities are sorted accordingly. Symmetric sampling is then applied after cascade matching, utilizing 128-bit descriptors. Isotropic and anisotropic filtering techniques are applied to the maximum filter response of the grayscale image. To normalize the descriptors, L2 normalization is performed on the RGB query image.

The resulting feature vectors are spatially organized, and Principal Component Analysis (PCA) is applied to reduce their dimensionality. To improve search efficiency, indexing and searching are performed based on a visual words representation of the database of visual features. The proposed method was evaluated using two datasets, Caltech-256 and Corel-1000, and compared to the standard HOG detector and descriptor. Experimental results show significant improvements in both average precision and average recall for the proposed method.

INTRODUCTION

The rapid expansion of multimedia databases and the increasing availability of high-bandwidth communication have led to the development of Content-Based Image Retrieval (CBIR) systems. CBIR allows users to search for images based on their visual content rather than relying on metadata or text-based tags. One of the primary challenges in designing a CBIR system is determining the appropriate image features for indexing and retrieval. In large image databases, retrieval based on content involves searching for images by comparing their

visual features, such as shape, texture, and color. CBIR can be broken down into two key stages:

1. **Indexing Stage**: Images in the database are stored in an organized manner. This phase involves extracting and cataloging image features (e.g., color, texture, shape) and storing them in an index, which links the features to the respective image in the database.

2. **Retrieval Stage**: When a user queries the system, images are retrieved based on a similarity measure between the query and the images in the database. The features of the query image are compared with those in the indexed database to retrieve the most similar results.

In traditional image retrieval methods, **Text-Based Image Retrieval (TBIR)** is used, where images are indexed using manually assigned keywords or tags. Search engines like Google and AltaVista employ TBIR by searching for text surrounding images, such as labels or captions. However, this approach has significant limitations. It often leads to difficulties in handling large image datasets and may not accurately capture the visual content of images. Different users can interpret the same image in various ways, leading to inconsistencies in search results.

In the 1990s, CBIR systems were introduced to overcome these limitations by using visual content to index and retrieve images. Unlike traditional text-based systems, CBIR directly analyzes and extracts features from the image itself. The key difference lies in how images are indexed and retrieved: CBIR systems focus on visual characteristics, while traditional systems rely on textual descriptions or keywords. However, retrieval results in CBIR are often not perfect matches but rather images that share visual similarities with the query.

Unlike traditional search methods that require exact matches, CBIR employs similarity measures, allowing for more flexible and efficient searches. Users may have different interests or perspectives on an image, and CBIR systems can retrieve various parts of the image that are most relevant to their query.

In terms of image search, manual keyword assignment is still common, but human perception of image content can vary, and textual descriptions are often inadequate or subjective. For instance, text descriptions cannot always accurately convey the texture or color of an image, and synonyms in image annotations can cause searches to miss relevant results. Furthermore, some images, such as those captured by surveillance cameras, may lack descriptive metadata entirely.

To address these limitations, CBIR systems automatically extract image features—such as color, texture, and spatial relationships between objects—and store them in a structured database for retrieval. The main goal of CBIR is to improve the efficiency of image indexing and retrieval by relying on the image's intrinsic properties rather than external annotations. One of the primary tasks of CBIR is feature extraction, where characteristics like pixel values are compared to quantify image similarity. These comparisons allow the system to retrieve visually similar images based on computed differences between their features.

Among the various visual properties used in CBIR, color plays a crucial role because it remains consistent across different scales and transformations. Despite this, human perception of visual content may still vary, and different users may find different parts of the image relevant. Nonetheless, CBIR continues to evolve as a powerful tool for efficient and content-based image retrieval.

The rapid expansion of multimedia databases and the increasing availability of high-bandwidth communication have led to the development of Content-Based Image Retrieval (CBIR) systems. CBIR allows users to search for images based on their visual content rather than relying on metadata or text-based tags. One of the primary challenges in designing a CBIR system is determining the appropriate image features for indexing and retrieval. In large image databases, retrieval based on content involves searching for images by comparing their visual features, such as shape, texture, and color. CBIR can be broken down into two key stages:

1. **Indexing Stage**: Images in the database are stored in an organized manner. This phase involves extracting and cataloging image features (e.g., color, texture, shape) and storing them in an index, which links the features to the respective image in the database.

2. **Retrieval Stage**: When a user queries the system, images are retrieved based on a similarity measure between the query and the images in the database. The features of the query image are compared with those in the indexed database to retrieve the most similar results.

In traditional image retrieval methods, **Text-Based Image Retrieval (TBIR)** is used, where images are indexed using manually assigned keywords or tags. Search engines like Google and AltaVista employ TBIR by searching for text surrounding images, such as labels or captions. However, this approach has significant limitations. It often leads to difficulties in handling large image datasets and may not accurately capture the visual content of images. Different users can interpret the same image in various ways, leading to inconsistencies in search results.

In the 1990s, CBIR systems were introduced to overcome these limitations by using visual content to index and retrieve images. Unlike traditional text-based systems, CBIR directly analyzes and extracts features from the image itself. The key difference lies in how images are indexed and retrieved: CBIR systems focus on visual characteristics, while traditional systems rely on textual descriptions or keywords. However, retrieval results in CBIR are often not perfect matches but rather images that share visual similarities with the query.

Unlike traditional search methods that require exact matches, CBIR employs similarity measures, allowing for more flexible and efficient searches. Users may have different interests or perspectives on an image, and CBIR systems can retrieve various parts of the image that are most relevant to their query.

In terms of image search, manual keyword assignment is still common, but human perception of image content can vary, and textual descriptions are often inadequate or subjective. For instance, text descriptions cannot always accurately convey the texture or color of an image, and synonyms in image annotations can cause searches to miss relevant results. Furthermore, some images, such as those captured by surveillance cameras, may lack descriptive metadata entirely.

To address these limitations, CBIR systems automatically extract image features—such as color, texture, and spatial relationships between objects—and store them in a structured database for retrieval. The main goal of CBIR is to improve the efficiency of image indexing and retrieval by relying on the image's intrinsic properties rather than external annotations. One of the primary tasks of CBIR is feature extraction, where characteristics like pixel values are compared to quantify image similarity. These comparisons allow the system to retrieve visually similar images based on computed differences between their features.

Among the various visual properties used in CBIR, color plays a crucial role because it remains consistent across different scales and transformations. Despite this, human perception of visual content may still vary, and different users may find different parts of the image relevant. Nonetheless, CBIR continues to evolve as a powerful tool for efficient and content-based image retrieval.

1. Related work

Content-Based Image Retrieval (CBIR) research can be classified into two main categories based on the types of features used for image retrieval: shape, texture, color, and region-based features. The current trend in CBIR focuses on extracting visual image features for more effective retrieval.

In [17], a study on content-based image retrieval using color histograms is presented. The authors used color histograms to retrieve similar images and implemented techniques to accelerate retrieval, such as projecting images and employing precise image matching. To improve search speed, they indexed the images in vector space, which enhanced the retrieval performance.

In [18], an Interactive Genetic Algorithm (IGA) was proposed to minimize the semantic gap between the retrieval results and user expectations. The authors utilized the HSV (Hue, Saturation, Value) color space, which is more aligned with human color perception by separating luminance from chrominance components. They also explored recurrence matrices and border histograms for better image retrieval performance.

Another study in [19] introduced a clustering technique combined with image mining to improve retrieval speed. The approach focused on the RGB color space, utilizing fuzzy C-means clustering to optimize image categorization. This method is aimed at reducing data loss during image retrieval.

A module descriptor in CBIR can significantly influence retrieval accuracy because it serves as a key feature for understanding objects in an image. Various extraction methods, such as mold-based and region-based shape extraction, are used. The mold-based method focuses on the outer boundary, while the region-based method considers the entire image area [20]. Additionally, wavelet decomposition has been applied in some CBIR systems to enhance performance, with fuzzy series fitting used to filter out irrelevant images. After filtering, the remaining images are compared to the query image to retrieve the most relevant results.

Texture, though not precisely defined, plays an important role in human image interpretation. It measures properties like roughness, smoothness, and symmetry. Texture can be analyzed using three primary methods: structural, statistical, and spectral. The structural method analyzes small patterns within the image, the statistical method uses statistical measures, and the spectral method applies frequency-domain filters. However, the structural method has limitations, especially in the context of natural systems that include irregular shapes [21, 22]. Furthermore, texture-based image retrieval has faced challenges, particularly with scaled or structured images, resulting in suboptimal performance due to high computational complexity.

In [23], the authors compared the performance of Gabor wavelets with traditional orthogonal wavelet functions for large structured image datasets. Gabor wavelets were found to be more effective in terms of overall performance. However, traditional orthogonal wavelets are not well suited for image decomposition in CBIR systems.

A CBIR system discussed in [24] extracts both structural and chromatic features from images using color histograms and Gabor filters. The system combines these features in linear combinations for image retrieval. This approach is computationally intensive, but it has demonstrated good results in retrieving images based on texture, color, and structural features.

In [25], a system was proposed for image retrieval based on color histograms. This method improves retrieval performance under transformations like rotation and flipping. Another technique discussed in [26] employs color distribution and pixel difference analysis for image retrieval. This approach, based on the K-means clustering algorithm, showed strong results across multiple image databases. Importantly, it remains unaffected by image displacement and rotation, but real-time analysis in dynamic environments still poses challenges.

In [27], a novel approach was presented that combines color histograms, texture consistency, and invariant moments to retrieve images based on edge regularity. The method was evaluated using precision and recall metrics, but environmental factors such as fog, snow, or precipitation were found to degrade surveillance image quality, impacting retrieval accuracy.

A CBIR model introduced in [28] combines Gabor filters and 3D histograms for feature extraction. Gabor filters capture texture features, while the 3D histogram is used to extract color features. The system leverages genetic algorithms to fine-tune the process. However, traditional methods often suffer from the "curse of dimensionality," leading to increased computational requirements and decreased performance. The efficiency of a CBIR system is often limited by the image descriptors used for retrieval, such as distance metrics, color descriptors, texture consistency, and shape descriptors [29].

2. Method & material

This study presents an efficient image retrieval system based on the Bag-of-Words (BOW) model. Initially, key regions of interest in each image from the database are identified using an affine-invariant detector. A fast binary descriptor is then applied to establish correspondences between points in the images. Descriptors are extracted from regions of varying sizes and subsequently reduced in dimensionality using Principal Component Analysis (PCA) [30]. To generate a feature vector, the query image is spatially organized to capture color-based features. A visual vocabulary is constructed by applying the k-means clustering algorithm to the combined features from the extracted descriptors and the color feature vector. This visual vocabulary is then used to assess the similarity between the feature vector of the query image and the images in the dataset.

2.1. Interest Region Detection

Interest regions are segments of an image that are distinct from their neighboring areas. The proposed detector identifies a combination of connected components formed by extremal regions at various threshold levels in the input image. These extremal regions, which remain stable at certain threshold values, are recognized as interest regions based on the extremal properties of the region intensity function, along with its outer boundary consisting of either darker or brighter pixels. Pixels above or equal to the threshold are classified as "white," while those below the threshold are classified as "black." The segmented regions are approximated using shapes such as parallelograms or ellipses. The parameters of these fitted shapes are then used to mathematically represent the interest regions identified in the grayscale image.

Let $t = 1, \ldots, m-1, m, m+1, \ldots$, denote the tested thresholds for a given grayscale image $\mathbb{I}$, where $\mathcal{R}1, \ldots, \mathcal{R}m-1, \mathcal{R}m, \mathcal{R}m+1$ are the nested extremal regions. The stability of these regions is expressed as a threshold function m [31]:

$$\Upsilon(m) = |\mathcal{R}m| \,/\, |\mathcal{R}m + \nabla\mathcal{R}m - \nabla| \quad (1)$$

In Equation (1), the boundary thickness is determined by the parameter ∇, and the cardinality is represented by $|\ .\ |$. The proposed detector is denoted as $\mathcal{R}m*$ with $\Upsilon m*$, which identifies stable regions by detecting threshold changes across a broad range. The combination of all extremal connected regions, denoted as $\boldsymbol{\delta}$, is acquired by thresholding and exhibits several important properties. First, $\boldsymbol{\delta}$ remains invariant under monotonic changes in image intensity, as it is solely based on the ordering of pixel intensities, and this order is preserved under monotonic transformations. Local affine or linear photometric changes do not affect $\boldsymbol{\delta}$. On the other hand, constant geometric transformations, such as pixel shifts between connected components, maintain the topology. As a result, after geometric transformations like non-linear continuous warping, homograph, or affine transformations, a transformed set $\boldsymbol{\delta}`$ of matching extremal regions is obtained. Thus, the combination of regions remains consistent under a wide range of photometric and geometric transformations, preserving the same cardinality. Since the extremal regions are fewer than the total number of image pixels, this approach is computationally efficient.

The enumeration of the set of extremal regions, $\boldsymbol{\delta}$, is performed in an efficient manner with a time complexity of $O(n \log\log n)$, where n represents the total number of pixels in the image. The process begins by sorting the image pixels based on intensity using the BINSORT algorithm [32], which has a linear complexity of $O(n)$ for small intensity values, such as {0, … , 255}. After sorting, pixels are arranged in ascending or descending order. The union-find algorithm [33] is then applied to maintain a list that tracks merging and growing connected components and their corresponding areas. The union-find algorithm operates efficiently with a complexity of $O(n \log\log n)$. During the enumeration process, a data structure is built to store each connected component's area as an intensity function. When two components are merged, the smaller one is incorporated into the larger one, eliminating any smaller components in the process.

2.2. Feature Description

After detecting the interest regions at a specific image location $\ell(x, y)$ with orientation $\emptyset$ and scale $\mathbb{s}$, a discriminative matching process is performed. To facilitate this, the image content and its local structure around ℓ need to be encoded into a descriptor that is both scale-invariant (relative to $\mathbb{s}$) and orientation-aligned (with $\emptyset$). The proposed method introduces a fast binary descriptor for this feature encoding. Image intensities are compared using a sampling pattern similar to the human retina, which is then used to generate a binary string through a cascading process.

2.3. Sampling Pattern

To perform comparisons across various sampling grids, pairs of pixel intensities are analyzed. Methods like ORB and BRIEF rely on randomly selected pixel pairs, while BRISK adopts a circular sampling pattern where points are evenly spaced along concentric rings, similar to the DAISY approach [34]. In contrast, the method proposed here utilizes a retinal-inspired sampling strategy. This circular pattern features a higher concentration of sampling points near the center, with point density decreasing exponentially toward the periphery, as illustrated in

Figure 1. Each sampling point is smoothed to reduce noise sensitivity.

While ORB and BRIEF apply kernels to all pixels within a patch, the retinal model follows BRISK in using a fixed kernel size for each sampling point. However, the proposed descriptor differs from BRISK in that the kernel size changes exponentially based on overlapping interest regions. Figure 4.1 demonstrates that each circle corresponds to a standard deviation, where Gaussian kernels are applied to sample points arranged in parallel. This log-polar distribution of Gaussian kernel sizes, inspired by the retinal structure, has shown to significantly enhance performance. Moreover, overlapping interest regions further improve descriptor effectiveness by increasing discriminative power through redundancy.

Let us now consider intensity values $\mathbb{I}i$ sampled at interest regions $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$, where:

$\mathbb{I}\mathcal{X} > \mathbb{I}\mathcal{Y}$, $\mathbb{I}\mathcal{Y} > \mathbb{I}\mathcal{Z}$, and $\mathbb{I}\mathcal{X} > \mathbb{I}\mathcal{Z}$ (2)

If the sampled regions do not overlap, the third condition ($\mathbb{I}\mathcal{X} > \mathbb{I}\mathcal{Z}$) provides no additional discriminative information. However, when regions overlap, this comparison may capture new and useful distinctions. Generally, incorporating redundancy allows the use of less sensitive sampling fields while still enhancing feature discrimination—an effect also observed in the overlapping receptive fields of the human retina [35].

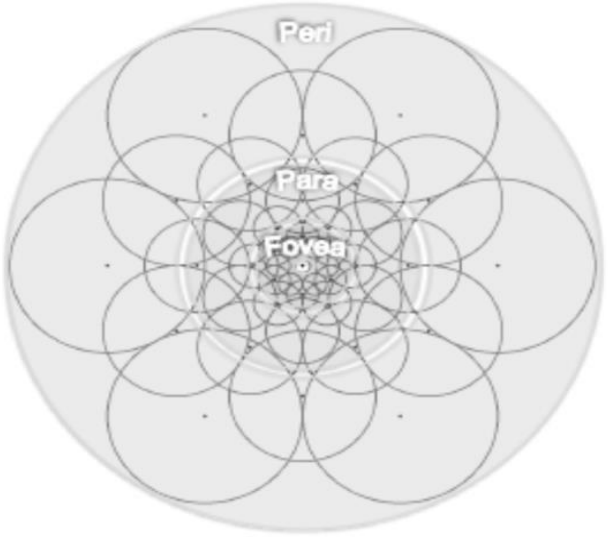

Figure 1: Representation of the sampling pattern

2.3.1. Coarse to fine descriptor

The proposed binary descriptor is built by applying a thresholding function to intensity differences between pairs of interest regions, each associated with a Gaussian kernel. As a result, a binary string EEE is generated using Difference of Gaussian (DoG), where each bit in the string is derived from intensity comparisons, as described in Equation (3) [36]:

$$E = \sum_{0 \le a < N1} 2aT(qa \quad)$$

In this equation, N represents the total length of the descriptor, and q_a denotes a specific pair of receptive fields. The thresholding function $T(q_a)$ is defined as:

$$T(q_a) = \begin{cases} 1 & \text{if } I(q_r1) - I(q_r2) > 0, \\ 0 & \text{otherwise} \end{cases}$$

(4)

As shown in Equation (4) [37], $I(q_r2)$ correspond, to the smoothed intensity of the second region in the receptive pair q_a. Since the number of potential pairs across multiple interest regions can easily reach thousands, many of them may contribute little to discriminative power. Thus, to enhance efficiency, pairs are selected based on their spatial distance and correlation.

To identify the most informative and uncorrelated pairs, a selection algorithm inspired by ORB is applied to training data. The steps are as follows:

1. A matrix M is created from approximately 50,000 extracted key points, where each row corresponds to a key point. Around 1,000 pairs are formed from 43 interest regions.
2. The mean value of each column is calculated. Pairs with a mean near 0.5 tend to exhibit high variance and are considered the most discriminative.
3. Columns are ranked based on variance.
4. The best columns (those with mean close to 0.5) are iteratively selected, ensuring low correlation between chosen features.

The selected DoG pairs follow a coarse-to-fine order, which is automatically determined. As illustrated in Figure 2, these pairs are grouped into four clusters, each containing 128 pairs. The first 512 pairs are used, as adding more does not significantly improve performance. The initial clusters mostly contain peripheral interest regions, while the final clusters are centered around the most focused areas—an arrangement resembling the human visual system. Peripheral regions help in object localization, whereas verification is refined using the densely packed foveal receptive fields.

Ultimately, the proposed feature descriptor is heuristic in nature and modeled after the human retina. It effectively captures and matches object features through a biologically inspired approach.

Equation (4) [37] represents the smoothed intensity $I(q^{r^2})$ of primary interested region of the pair q_a. With some dozens of interested regions, a large descriptor can lead thousands of pairs. Whereas, an image cannot be described efficiently by most of the pairs. Therefore, pairs are selected based on spatial distance. Thus, the chosen pairs are non-discriminant and highly correlated. Therefore, an algorithm same as ORB is applied on training data to locate the best pairs. The algorithm is as follows:

1) A matrix $\mathcal{M}$ is created consisting of almost 50,000 extracted key points. Every row related to key point. Almost 1,000 pairs are used with 43 interested regions.
2) Calculate the mean value of every column. The discriminant feature is obtained with high variance. The mean value of 0.5 results in highest binary distribution variance.
3) The columns are sorted according to the maximum variance.
4) Remaining columns are added recursively with low correlation while keeping the best column with 0.5 as mean value.

DoG (the pairs) coarse-to-fine ordering is selected automatically. Figure 2 suggests the selected pairs grouped into 4 clusters (128 pairs in one group). First 512 pairs considered as relevant whereas if the pairs are increased, the performance will not increase. A symmetrical sample is detected with the global gradient. Peripheral interested regions are involved in initial cluster while the highly centered regions are implicated in last ones. This seems as a reminiscent for human eye behavior. The interested object is located by using perifoveal interested regions. Therefore, the verification is performed by maximum densely distributed area of fovea receptive fields. However, the proposed feature descriptor is heuristic which is based on human retina model, used to detect and match object.

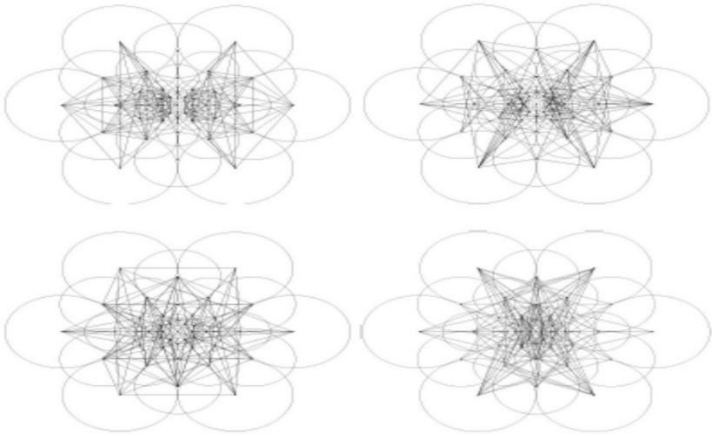

Figure 2: Representation of the analysis from coarse-to-fine

2.3.2. Saccadic Search

Human vision operates through rapid, irregular eye movements known as saccades, which allow the eyes to continuously scan a scene. These movements stem from the structure of retinal cells. As previously discussed, the fovea—the region responsible for high-acuity vision—requires detailed information to accurately identify and compare objects. In contrast, the perifoveal areas, which detect lower-frequency information, are primarily responsible for forming an initial approximation of where objects of interest might be located within the visual field.

Inspired by this biological mechanism, the research introduces a method that mimics saccadic processing by analyzing feature descriptors in multiple stages. Initially, a raw 16-byte feature descriptor is employed. This form of descriptor is widely adopted in the literature, with over 90% of existing works utilizing the 16-byte format due to its compatibility with hardware constraints. Specifically, it aligns with Single Instruction, Multiple Data (SIMD) operations on Intel processors, enabling efficient parallel processing.

The approach involves a cascading series of comparisons to speed up the matching and adaptation phase, allowing for rapid identification of similar objects across images. In the first stage, only a portion of the 16-byte descriptor is evaluated to quickly filter out irrelevant matches. For keypoint rotation estimation, local gradient information can be incorporated using selected BRISK-style point pairs [38]. The connection between a keypoint's orientation and its position relative to the global image center is illustrated in Figure 3.

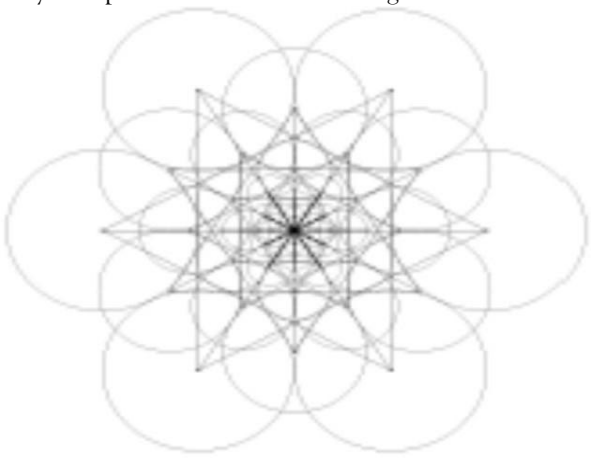

Figure 3: Representation of the selected pairs of orientation.

In Equation (5), **M** denotes the number of pairwise combinations in **G**, and ***Pr*1** refers to the two-dimensional spatial vector representing the center of the receptive field. Instead of utilizing several hundred BRISK point pairs, this method selects 45 optimized pairs for greater efficiency.

The proposed detector focuses on identifying intensity-based features such as corners and edges within grayscale images. It begins by detecting pixels within a specific intensity range, forming connected regions based on those values. Once bright and dark regions are identified, pixel intensities are sorted accordingly.

Following this, symmetric sampling is carried out after performing cascade matching with a 128-bit binary descriptor. The grayscale image is further processed using both isotropic and anisotropic filtering on the output of a collapsed maximum filter response.

To reduce the dimensionality of the resulting feature vectors, Principal Component Analysis (PCA) is applied, as illustrated in Figure 4. This step ensures a compact and discriminative representation for efficient image analysis and matching.

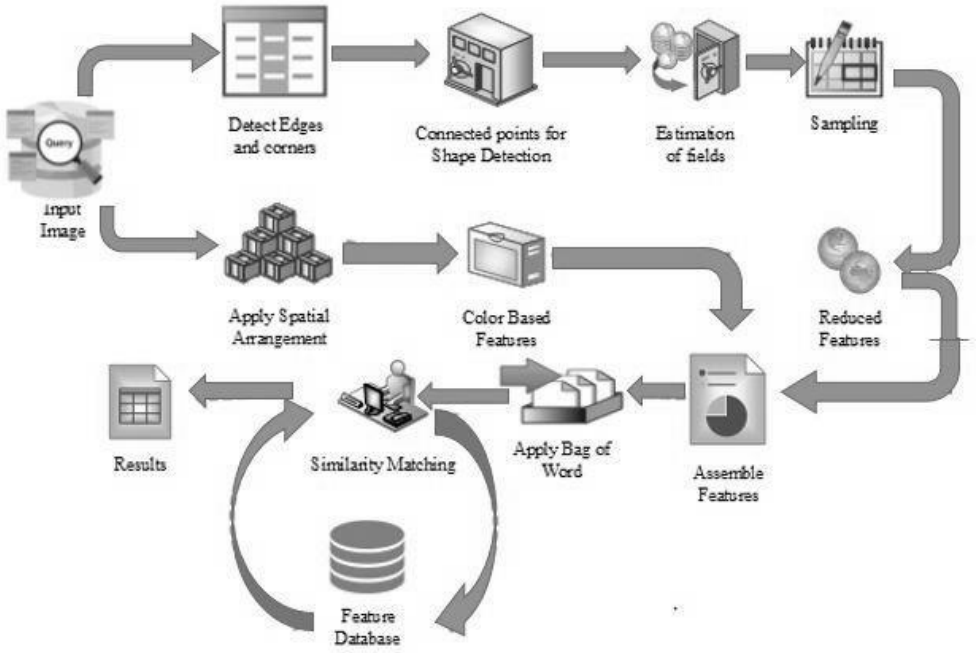

Figure 4: Proposed Method

2.4. Feature Vector Reduction

A linear dimensionality reduction technique, Principal Component Analysis (PCA), is employed to compress the descriptor by retaining the top 128 eigenvalues. PCA uses an orthogonal matrix to project the original feature vectors onto a new set of axes. These axes, known as principal components, are derived from the corresponding eigenvectors of the data's covariance matrix and represent uncorrelated directions of maximum variance.

The principal components are primarily influenced by Gaussian-distributed features. To minimize the descriptor size while preserving essential information, only the leading eigenvectors—those associated with the highest eigenvalues—are selected. These eigenvectors are ordered in descending magnitude based on their corresponding eigenvalues, ensuring that the most significant components are retained.

2.5. Spatial Arrangement

Spatial structuring is applied to the RGB query image to preserve the arrangement of features. These feature locations are determined using estimated geometric transformations. While color histograms are commonly used to capture the color distribution, they lack the ability to represent spatial relationships. To address this limitation, a distance metric is introduced to account for spatial correlations in color variations.

Let $\mathcal{Q}$ denote the query image, where the quantized color values are represented as $\mathcal{W}_1$, $\mathcal{W}_2$, ..., $\mathcal{W}\mathcal{k}$ for $\mathcal{k}$ different color bins. For any pixel position $\mathcal{d}$ = ($\mathcal{u}$, $\mathcal{v}$) within the image $\mathcal{Q}$, the color information at that location is considered for analysis [119].

The spatial color distance between two-pixel locations, $\mathcal{d}_1$ = ($\mathcal{u}_1$, $\mathcal{v}_1$) and $\mathcal{d}_2$ = ($\mathcal{u}_2$, $\mathcal{v}_2$), is calculated using the following metric [119].

$$|d1 - d2| \approx \max(|u1 - u2|, |v1 - v2|)$$

This method allows the system to not only measure color similarity but also account for the spatial placement of colors within the image. The resulting histogram $\mathcal{S}$ for the image $\mathcal{Q}$ reflects both color presence and its approximate spatial distribution [119].

$|d1-d2| \approx \max\{|u_1 - u_2|, |v_1 - v_2|\}$

The histogram S corresponding to image Q is presented below, as referenced in [119].

$$S_{T_m}(Q) \triangleq r^2 \underset{e \in Q}{d}[e \in Q \quad T_m]$$

The likelihood of a pixel having the color Tm in image Q is calculated using the expression $ST_m(Q)/r$, where Tm represents a specific color value of a pixel in the image. The histogram is generated with linear time complexity, denoted as O(n). Instead of utilizing a correlogram, a fixed spatial distance T is used for the image Q as described in [119].

$$C^f(Q) \approx \varepsilon 1\ [d1,d2 \in Q \| d1-d2\| = f]$$

Equation (6) outlines the spatial configuration of image elements. The probability C represents the likelihood of a distance f between a specific color pixel TU and a reference pixel. Equation (7) further illustrates the spatial correlation among pixels that share identical color values.:

$$O^f_T(Q) \triangleq C^f_{T,T}(Q)$$

In Equation (7), the probability O represents the likelihood of a color pixel T occurring at a distance f. Once spatial structuring of the image is completed, a feature vector based on color is extracted. This color feature vector is then combined with the dimensionally reduced feature vector and passed into the Bag of Words (BoW) framework. Subsequently, indexing and search operations are carried out using the visual word representations derived from the feature database to retrieve relevant results.

3.5 BOW

To store the extracted feature vectors and assess similarity between key points, a Bag-of-Words (BoW) model is employed. This model generates a codebook that quantizes key-point vectors and assigns labels to important image regions using representative visual words. The codebook is created by quantizing geometric and content features. For this purpose, the widely-used and scalable K-means clustering algorithm is applied to cluster and classify the image features. Once the codebook is generated, each image is represented by a histogram that counts the frequency of visual words. To normalize these histograms, the Term Frequency-Inverse Document Frequency (tf-idf) method is applied. Prior to clustering, the histograms are normalized and features are refined by filtering out repeated, rare, or irrelevant visual words.

3.6 Similarity Matching

Similarity between the feature database and a query image is calculated based on the extracted features. For effective retrieval of relevant images, Euclidean distance is used to measure similarity.

3. Results & Implementation

This section presents the evaluation results of the proposed approach and compares its performance against other descriptors based on specific criteria.

4.1 Result Analysis

Initially, the color images are converted into grayscale to facilitate the detection of intensity-based local interest points. Using optimized sliding windows, global features around these points of interest are extracted. The extracted features are then combined through a proposed feature reformulation technique, which concatenates the feature vectors and generates coefficients for the restructured data. This processed data is subsequently classified using a Support Vector Machine (SVM), which operates in two phases: training and testing. Accuracy, representing the proportion of correctly predicted positive samples, and recall, indicating the true positive rate, are computed for datasets of various sizes, including those with 100 images. The performance, robustness, and efficiency of the algorithm are evaluated by applying it across

different categories within each dataset. The formulas for calculating precision and recall rates are provided below:

$$Precision = \frac{\text{relevant images + retrieved images}}{retrieved\ images}$$

$$Recall = \frac{\text{relevant images+retrieved images}}{relevant\ images}$$

4.2 Dataset Detail

Selecting the appropriate image database is a crucial step in designing an effective image search system. In our experiments, each image in the database is used as a query image by applying rotations. For each query, the system retrieves the most relevant images from the database. The similarity between database images and the query image is evaluated using various distance metrics. If the retrieved image matches the query image, the system is considered to have successfully located the target; otherwise, the desired image is not recovered. This study focuses on assessing the performance of feature generalization techniques using two image datasets: Caltech-256 and Corel-1000.[30].

4.2.1 Corel-1000

The Corel-1000 dataset contains a total of 1,000 images divided into 10 categories, with each category consisting of 100 images. The dimensions of each image are either 384x256 or 256x384 pixels. The categories include African people and villages, beaches, homes, buses, dinosaurs, elephants, flowers, horses, mountains, and food. To evaluate relevancy, a query image is provided, and images retrieved that match the query are considered relevant, while those that do not match are deemed irrelevant.

Figure 5: One sample image from each category Corel-1000 Dataset

4.2.2 Caltech-256

The Caltech-256 dataset consists of 30,607 images collected through Google Image Search. These images are categorized into 257 different classes based on their content and relevance. Compared to Caltech-101, classification within Caltech-256 is more challenging due to greater variability among the images. For our experiments, we focused on 15 specific categories including teapot, tomato, billiards, backpack, bulldozer, bowling ball, teddy bear, boxing gloves, bonsai, airplane, butterfly, clock, spider, cactus, and swan. From these categories, a subset of

1,500 images was used, with each category containing 100 images. Some categories are difficult to classify due to complex patterns, while others are challenging because of their foreground elements or background objects. The experiments were conducted using these selected images

Figure 6: One Sample image of each category Caltech-256 Dataset.

4.3. Retrieved images of Datasets

In the Corel-1000 dataset, query images were selected at random, and the system retrieved images that were either relevant or irrelevant as results. As shown in Figure 7, all retrieved images correspond closely to the given query image, demonstrating high accuracy. For each query, 20 images were retrieved, all matching the category of the query image. For example, when a horse image was used as the query, all 20 retrieved images were also related to horses, confirming the effectiveness of the feature extraction process based on category matching. This outcome highlights the strong performance of the proposed algorithm.

Compared to the HOG descriptor, the proposed method delivers superior retrieval results. Figures 8 and 9 further illustrate the outstanding retrieval accuracy of the Corel-1000 dataset using our approach, with all returned images closely related to their respective queries.

Query Image

Figure 7: 1st Retrieved images of Corel-1000

Query Image

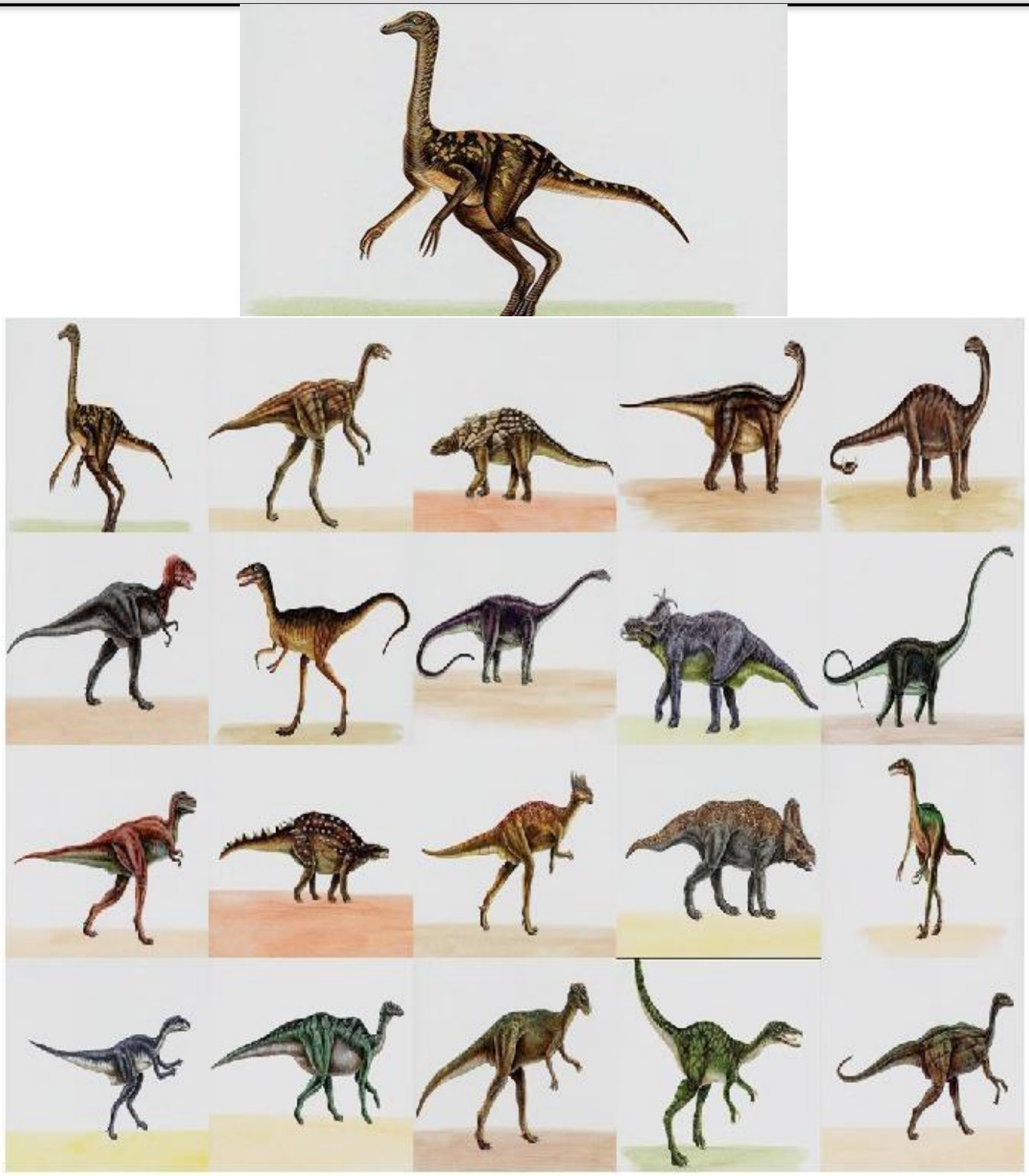

Figure 8: 2nd Retrieved images of Corel-1000

Figure 9 presents the results for the Corel-1000 dataset. When a query image is provided, the system retrieves 20 images, all of which are relevant to the query. The proposed algorithm demonstrates superior performance in retrieving these relevant images.

Figure 9: 1st Retrieved images of Caltech-256

4.3 Experimental results

The effectiveness and efficiency of the proposed method were evaluated by comparing it with the HOG descriptor. Two benchmark datasets—Corel-1000 and Caltech-256—were used for experimentation, with various categories selected from each to ensure comprehensive evaluation. The results demonstrate that the proposed algorithm performs exceptionally well across multiple image categories. To enhance computational efficiency, the color images were converted to grayscale, which simplifies processing by focusing on intensity values to detect key points. Optimized sliding windows were used to extract global features from these interest

points. Feature extraction plays a vital role in minimizing the data required to represent large sets of digital images, thereby reducing the computational burden Additionally, the algorithm efficiently handles complex data analysis and lowers the computational cost associated with classification. The proposed method achieves superior classification performance with a time complexity of O(n), outperforming other existing algorithms in terms of speed and accuracy.

4.3.1 Average Precision Performance for Datasets

Figure 10 illustrates a comparison of average precision rates for the Corel-1000 dataset, highlighting the performance differences between the proposed algorithm and HOG. Experiments were conducted on both the Corel-1000 and Caltech-256 datasets. The proposed method delivers strong results across multiple categories within the Corel-1000 dataset, including Africa, Beach, Building, Bus, Dinosaur, Elephant, Flowers, Horse, Mountains, and Food, achieving an average precision of at least 72%.

While HOG performs well in certain categories like Bus, the overall performance of the proposed approach is significantly better. The results clearly demonstrate the proposed algorithm's effectiveness and robustness across a wide range of image classes when compared to existing state-of-the-art methods. As shown in Figure 10, the proposed algorithm consistently outperforms HOG in most categories.

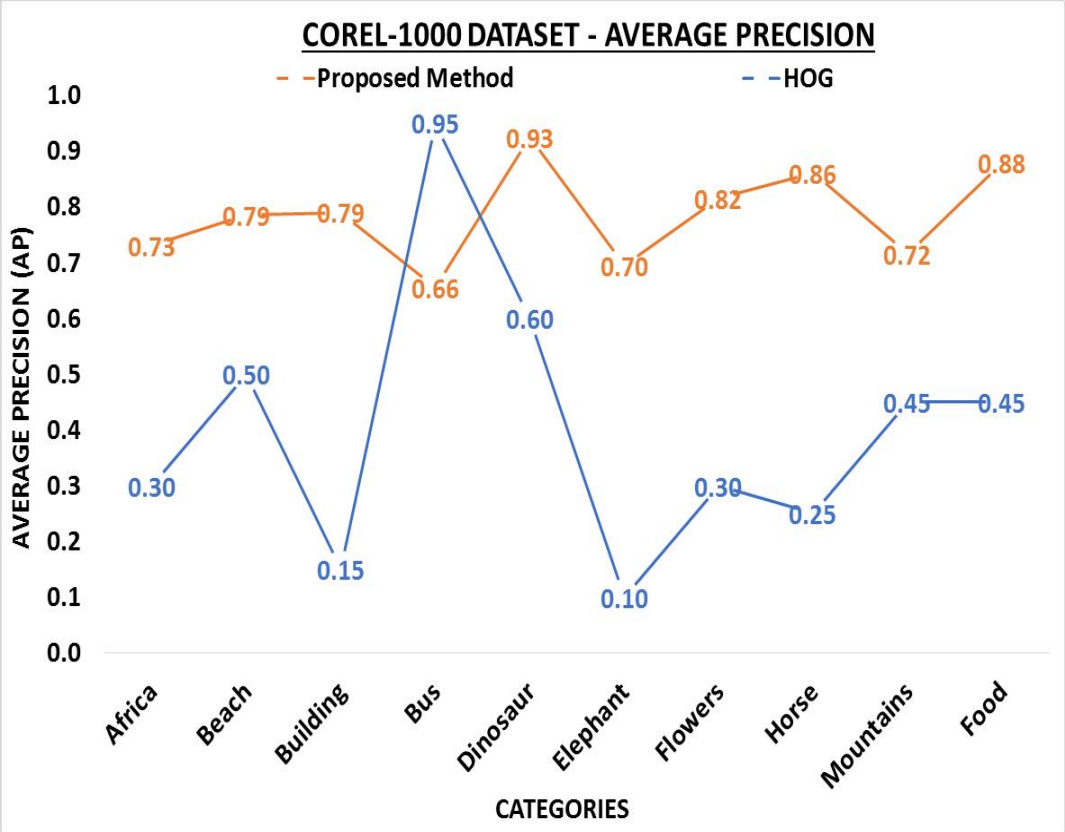

Figure 10: Average Precision for Corel-1000

As illustrated in Figure 11, several state-of-the-art algorithms perform poorly on the Caltech-256 dataset, with accuracy often falling below 50%. In contrast, the proposed algorithm achieves significantly higher average precision (AP), ranging from 65% to 75% across various categories. The comparative results clearly demonstrate the effectiveness of the proposed method, particularly in

categories such as Tomato, Backpack, Billiards, Bonsai, and Teapot, where it consistently outperforms competing approaches.

While HOG achieves relatively strong results in a few specific categories, the proposed algorithm delivers more consistent and superior performance overall. Its robustness and efficiency are especially evident in categories where other methods struggle to reach even 50% accuracy. The average precision achieved by the proposed method is approximately 75%, indicating a notable improvement over existing techniques.

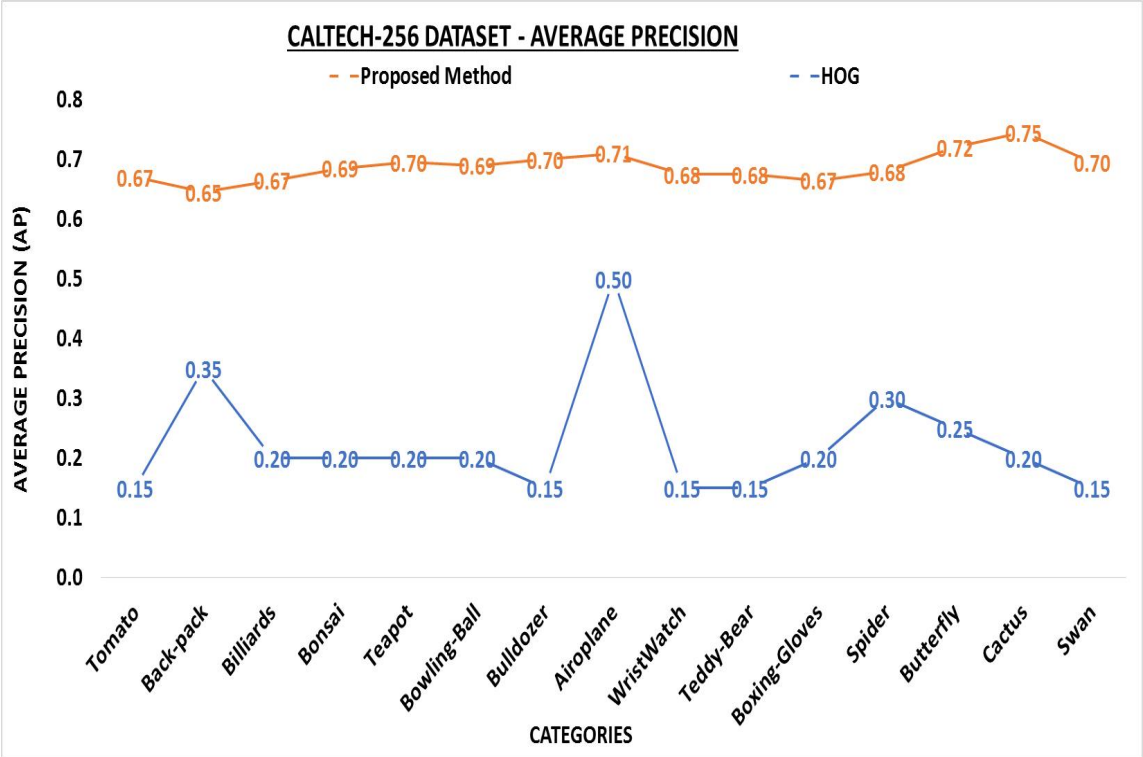

Figure 11: Average Precision for Caltech-256

The proposed algorithm demonstrates superior performance across the majority of categories within the Caltech-256 dataset when compared to alternative methods. While other algorithms often achieve average precision rates below 30% in many categories, the proposed method proves to be more robust and efficient. A comparative analysis of mean precision across various categories—such as Tomato, Backpack, Billiards, Bonsai, Teapot, Bowling Ball, Bulldozer, Airplane, Wristwatch, Teddy Bear, Boxing Gloves, Spider, Butterfly, Cactus, and Swan—highlights the consistently strong performance of the proposed approach. In contrast, the HOG algorithm records an average precision of less than 35% in several of these categories.

Average Precision Comparison on Caltech-256 Dataset

The table below highlights the average precision scores for various categories in the Caltech-256 dataset, comparing the performance of the proposed method with the traditional HOG descriptor. It is evident that the proposed method consistently outperforms HOG across all tested categories.

Table 1: Average Precision of Caltech -256

Average Precision for Caltech-256		
Category	**Proposed Method**	**HOG**
Tomato	0.67	0.15
Back-pack	0.65	0.35
Billiards	0.67	0.20
Bonsai	0.69	0.20
Teapot	0.70	0.20
Bowling Ball	0.69	0.20
Bulldozer	0.70	0.15
Airplane	0.71	0.50
Wristwatch	0.68	0.15
Teddy Bear	0.68	0.15
Boxing Gloves	0.67	0.20
Spider	0.68	0.30
Butterfly	0.72	0.25
Cactus	0.75	0.20

Average Precision Comparison on Corel-1000 Dataset

The following table summarizes the average precision values obtained using the proposed approach compared to the HOG descriptor across various categories in the Corel-1000 dataset. The proposed algorithm exhibits substantial performance improvements in most categories:

Table 2: Average Precision for Corel-1000

Average Precision for Corel-1000		
Category	**Proposed Method**	**HOG**
Africa	0.73	0.30
Beach	0.79	0.50
Building	0.79	0.15
Bus	0.66	0.95
Dinosaur	0.93	0.60
Elephant	0.70	0.10
Flowers	0.82	0.30
Horse	0.86	0.25
Mountains	0.72	0.45
Food	0.88	0.45

Tables 1 and 2 present the performance outcomes across all categories of the Caltech-256 and Corel-1000 datasets. The proposed method achieves strong results in nearly every category. While HOG demonstrates solid performance in specific classes, such as "Bus," the proposed algorithm outperforms it in several categories within the Corel-1000 dataset, including Africa, Beach, Buildings, Dinosaurs, Elephants, Flowers, Horses, Mountains, and Food.

Figure 12 illustrates a comparison of the average recall rates between the proposed method and HOG. Similarly, Figure 13 compares the average recall results for the Caltech-256 dataset, showing that the proposed algorithm consistently delivers improved performance across various categories. A lower repetition rate in retrieval corresponds to higher precision, while a higher repetition rate indicates reduced accuracy.

For the Corel-1000 dataset, performance comparisons are also provided in Figure 11. Although HOG performs well in certain classes, the proposed algorithm demonstrates superior overall accuracy, particularly in more complex and visually diverse categories.

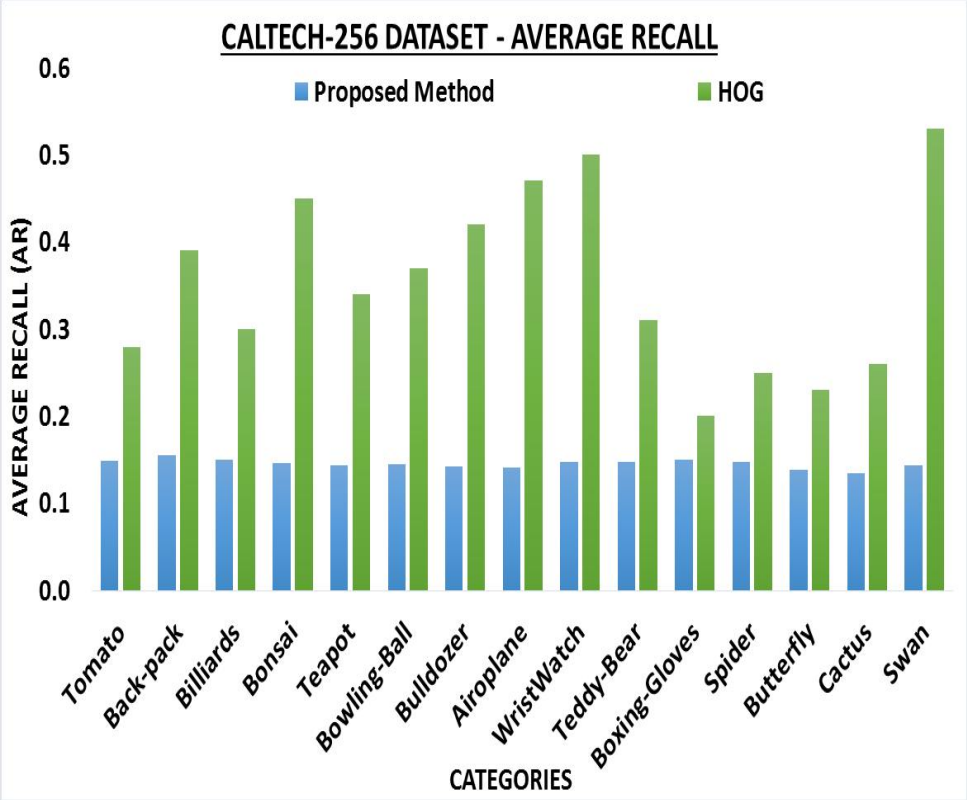

Figure 12: Average Recall for Caltech-256

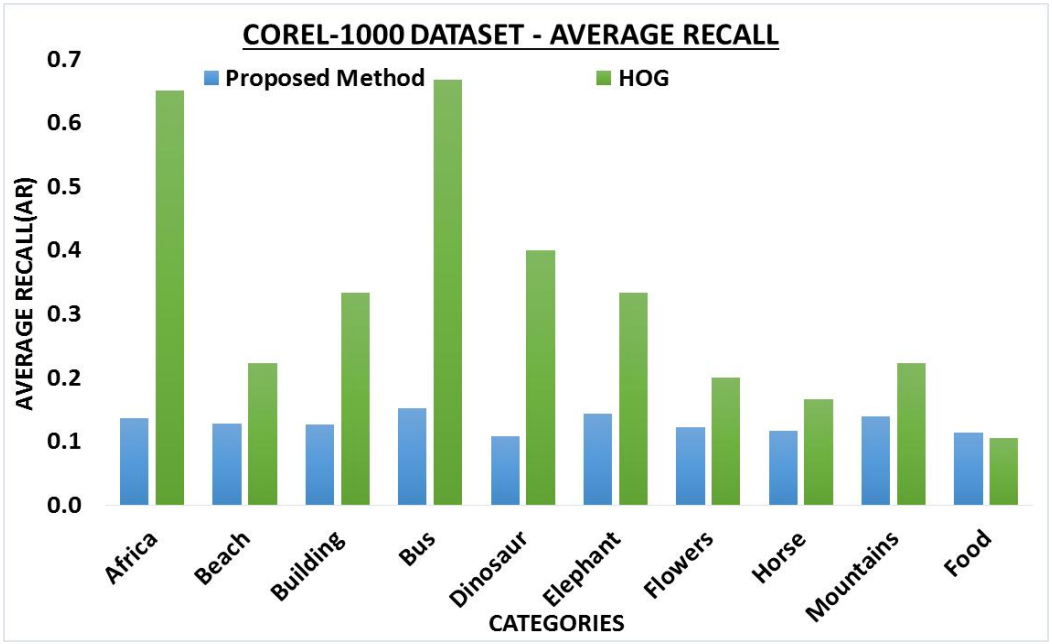

Figure 13: Average Recall for Corel-1000.

The proposed algorithm was evaluated using various benchmark datasets, including Caltech-256 and Corel-1000, to assess its accuracy, robustness, and overall efficiency. Figure 12 presents the Average Recall results obtained on the Corel-1000 dataset. These results demonstrate that the proposed method consistently outperforms traditional approaches, such as the Histogram of Oriented Gradients (HOG), across multiple categories.

Notably, the algorithm achieved superior recall performance in specific categories like Africa, Beach, Buildings, Dinosaurs, Elephants, Flowers, Horses, Mountains, and Food. The average recall rate ranges from 0.10 to 0.11, indicating reliable retrieval performance. Compared to HOG, the proposed method delivers higher recall values, highlighting its effectiveness in retrieving relevant images across diverse classes in the Corel-1000 dataset.

5 Conclusion

The rapid growth of multimedia databases, along with advancements in communication bandwidth and the increasing complexity of visual data, has driven the development of Content-Based Image Retrieval (CBIR). CBIR enables the retrieval of visually similar images by analyzing the actual content of images rather than relying on textual metadata. One of the key challenges in designing CBIR systems is selecting the most relevant visual features—such as color, texture, and shape—that effectively represent the image.

In large-scale image collections, retrieving relevant images based on visual similarity becomes a complex task. CBIR focuses on addressing this challenge by leveraging visual content features. A significant portion of computer vision applications relies on the matching of keypoints across images. Over the past few decades, several reliable and efficient algorithms have been introduced for keypoint detection. However, there is an increasing need for more compact, faster, and noise-robust descriptors that are invariant to rotation and scale.

This research introduces a novel algorithm that integrates a keypoint detector with a feature descriptor. The proposed detector extracts intensity-based corners and edges from grayscale images and forms connected regions based on the intensity range of detected pixels. After identifying bright and dark regions, pixel intensities are sorted. Symmetric sampling is conducted following a 128-bit cascade matching process. The grayscale image is then refined using both isotropic and anisotropic filtering applied to the output of a collapsed maximum filter. L2 normalization is applied to the RGB query image. The extracted feature vectors are spatially organized, and Principal Component Analysis (PCA) is employed to reduce the dimensionality of the feature set. To facilitate efficient retrieval, visual word indexing and search are conducted using a bag-of-visual-words representation from the feature database. The effectiveness of the proposed method was validated on two benchmark datasets: Caltech-256 and Corel-1000. The results were promising, with Caltech-256 achieving a mean average precision (mAP) of 0.65 and a mean recall of 0.18. Corel-1000 also demonstrated strong performance, with an mAP of 0.75 and a recall rate of 0.11. Additionally, the proposed method was compared with the Histogram of Oriented Gradients (HOG) descriptor and detector, showing superior results in both precision and recall metrics.

Future work will focus on the reconstruction of personalized 3D scenes from image projections and exploring more advanced similarity measures by leveraging statistical properties of visual data.

REFERENCE

Khan, S.U.R., Asif, S., Bilal, O. et al. Lead-cnn: lightweight enhanced dimension reduction convolutional neural network for brain tumor classification. Int. J. Mach. Learn. & Cyber. (2025). https://doi.org/10.1007/s13042-025-02637-6.

Khan, S.U.R., et al. Robust&Precise Knowledge Distillation-based Novel Context-Aware Predictor for Disease Detection in Brain and Gastrointestinal. 2025.

Hekmat, A., et al., Brain tumor diagnosis redefined: Leveraging image fusion for MRI enhancement classification. Biomedical Signal Processing and Control, 2025. 109: p. 108040.

Khan, Z., Hossain, M. Z., Mayumu, N., Yasmin, F., & Aziz, Y. (2024, November). Boosting the

Prediction of Brain Tumor Using Two Stage BiGait Architecture. In 2024 International Conference on Digital Image Computing: Techniques and Applications (DICTA) (pp. 411-418). IEEE.

Khan, S. U. R., Raza, A., Shahzad, I., & Ali, G. (2024). Enhancing concrete and pavement crack prediction through hierarchical feature integration with VGG16 and triple classifier ensemble. In 2024 Horizons of Information Technology and Engineering (HITE)(pp. 1-6). IEEE https://doi. org/10.1109/HITE63532.

Khan, S.U.R., Zhao, M. & Li, Y. Detection of MRI brain tumor using residual skip block based modified MobileNet model. Cluster Comput 28, 248 (2025). https://doi.org/10.1007/s10586-024-04940-3

Khan, U. S., & Khan, S. U. R. (2024). Boost diagnostic performance in retinal disease classification utilizing deep ensemble classifiers based on OCT. Multimedia Tools and Applications, 1-21.

Raza, A., & Meeran, M. T. (2019). Routine of encryption in cognitive radio network. Mehran University Research Journal of Engineering & Technology, 38(3), 609-618.

Waqas, M., Tahir, M. A., & Khan, S. A. (2023). Robust bag classification approach for multi-instance learning via subspace fuzzy clustering. Expert Systems with Applications, 214, 119113.

Al-Khasawneh, M. A., Raza, A., Khan, S. U. R., & Khan, Z. (2024). Stock Market Trend Prediction Using Deep Learning Approach. Computational Economics, 1-32.

Khan, U. S., Ishfaque, M., Khan, S. U. R., Xu, F., Chen, L., & Lei, Y. (2024). Comparative analysis of twelve transfer learning models for the prediction and crack detection in concrete dams, based on borehole images. Frontiers of Structural and Civil Engineering, 1-17.

Khan, S. U. R., & Asif, S. (2024). Oral cancer detection using feature-level fusion and novel self-attention mechanisms. Biomedical Signal Processing and Control, 95, 106437.

Farooq, M. U., Khan, S. U. R., & Beg, M. O. (2019, November). Melta: A method level energy estimation technique for android development. In 2019 International Conference on Innovative Computing (ICIC) (pp. 1-10). IEEE.

Raza, A.; Meeran, M.T.; Bilhaj, U. Enhancing Breast Cancer Detection through Thermal Imaging and Customized 2D CNN Classifiers. VFAST Trans. Softw. Eng. 2023, 11, 80–92.

Dai, Q., Ishfaque, M., Khan, S. U. R., Luo, Y. L., Lei, Y., Zhang, B., & Zhou, W. (2024). Image classification for sub-surface crack identification in concrete dam based on borehole CCTV images using deep dense hybrid model. Stochastic Environmental Research and Risk Assessment, 1-18.

Raza, A., Salahuddin, & Inzamam Shahzad. (2024). Residual Learning Model-Based Classification of COVID-19 Using Chest Radiographs. Spectrum of Engineering Sciences, 2(3), 367–396.

Khan, S.U.R.; Asif, S.; Bilal, O.; Ali, S. Deep hybrid model for Mpox disease diagnosis from skin lesion images. Int. J. Imaging Syst. Technol. 2024, 34, e23044.

Raza, A., Soomro, M. H., Shahzad, I., & Batool, S. (2024). Abstractive Text Summarization for Urdu Language. Journal of Computing & Biomedical Informatics, 7(02).

Khan, S.U.R.; Zhao, M.; Asif, S.; Chen, X.; Zhu, Y. GLNET: Global–local CNN's-based informed model for detection of breast cancer categories from histopathological slides. J. Supercomput. 2023, 80, 7316–7348.

Hekmat, Arash, Zuping Zhang, Saif Ur Rehman Khan, Ifza Shad, and Omair Bilal. "An attention-fused architecture for brain tumor diagnosis." Biomedical Signal Processing and Control 101 (2025): 107221.

Waqas, M., Tahir, M. A., & Qureshi, R. (2023). Deep Gaussian mixture model based instance relevance estimation for multiple instance learning applications. Applied intelligence, 53(9), 10310-10325.

Khan, S.U.R.; Zhao, M.; Asif, S.; Chen, X. Hybrid-NET: A fusion of DenseNet169 and advanced machine learning classifiers for enhanced brain tumor diagnosis. Int. J. Imaging Syst. Technol. 2024, 34, e22975.

M. Wajid, M. K. Abid, A. Asif Raza, M. Haroon, and A. Q. Mudasar, "Flood Prediction System Using IOT & Artificial Neural Network", VFAST trans. softw. eng., vol. 12, no. 1, pp. 210–224, Mar. 2024.

M. Waqas, Z. Khan, S. U. Ahmed and A. Raza, "MIL-Mixer: A Robust Bag Encoding Strategy for Multiple Instance Learning (MIL) using MLP-Mixer," 2023 18th International Conference on Emerging Technologies (ICET), Peshawar, Pakistan, 2023, pp. 22-26.

Khan, S.U.R.; Raza, A.;Waqas, M.; Zia, M.A.R. Efficient and Accurate Image Classification Via Spatial Pyramid Matching and SURF Sparse Coding. Lahore Garrison Univ. Res. J. Comput. Sci. Inf. Technol. 2023, 7, 10–23.

Shahzad, Inzamam, Asif Raza, and Muhammad Waqas. "Medical Image Retrieval using Hybrid Features and Advanced Computational Intelligence Techniques." Spectrum of engineering sciences 3, no. 1 (2025): 22-65.

Farooq, M.U.; Beg, M.O. Bigdata analysis of stack overflow for energy consumption of android framework. In Proceedings of the 2019 International Conference on Innovative Computing (ICIC), Lahore, Pakistan, 1–2 November 2019; pp. 1–9.

HUSSAIN, S., Raza, A., MEERAN, M. T., IJAZ, H. M., & JAMALI, S. (2020). Domain Ontology Based Similarity and Analysis in Higher Education. IEEEP New Horizons Journal, 102(1), 11-16.

Waqas, M., Tahir, M. A., Al-Maadeed, S., Bouridane, A., & Wu, J. (2024). Simultaneous instance pooling and bag representation selection approach for multiple-instance learning (MIL) using vision transformer. Neural Computing and Applications, 36(12), 6659-6680.

Shahzad, I., Khan, S. U. R., Waseem, A., Abideen, Z. U., & Liu, J. (2024). Enhancing ASD classification through hybrid attention-based learning of facial features. Signal, Image and Video Processing, 1-14.

Khan, S. R., Raza, A., Shahzad, I., & Ijaz, H. M. (2024). Deep transfer CNNs models performance evaluation using unbalanced histopathological breast cancer dataset. Lahore Garrison University Research Journal of Computer Science and Information Technology, 8(1).

Bilal, Omair, Asif Raza, and Ghazanfar Ali. "A Contemporary Secure Microservices Discovery Architecture with Service Tags for Smart City Infrastructures." VFAST Transactions on Software Engineering 12, no. 1 (2024): 79-92.

Asif Raza, Inzamam Shahzad, Ghazanfar Ali, and Muhammad Hanif Soomro. "Use Transfer Learning VGG16, Inception, and Reset50 to Classify IoT Challenge in Security Domain via Dataset Bench Mark." Journal of Innovative Computing and Emerging Technologies 5, no. 1 (2025).

Waqas, M., Ahmed, S. U., Tahir, M. A., Wu, J., & Qureshi, R. (2024). Exploring Multiple Instance Learning (MIL): A brief survey. Expert Systems with Applications, 123893.

Khan, S. U. R., Asif, S., Zhao, M., Zou, W., Li, Y., & Li, X. (2025). Optimized deep learning model for comprehensive medical image analysis across multiple modalities. Neurocomputing, 619, 129182.

Khan, S. U. R., Asif, S., Zhao, M., Zou, W., & Li, Y. (2025). Optimize brain tumor multiclass classification with manta ray foraging and improved residual block techniques. Multimedia Systems, 31(1), 1-27.

Khan, S. U. R., Asim, M. N., Vollmer, S., & Dengel, A. (2025). AI-Driven Diabetic Retinopathy Diagnosis Enhancement through Image Processing and Salp Swarm Algorithm-Optimized Ensemble Network. arXiv preprint arXiv:2503.14209.

Khan, Z., Khan, S. U. R., Bilal, O., Raza, A., & Ali, G. (2025, February). Optimizing Cervical Lesion Detection Using Deep Learning with Particle Swarm Optimization. In 2025 6th International Conference on Advancements in Computational Sciences (ICACS) (pp. 1-7). IEEE.

Khan, S.U.R., Raza, A., Shahzad, I., Khan, S. (2025). Subcellular Structures Classification in Fluorescence Microscopic Images. In: Arif, M., Jaffar, A., Geman, O. (eds) Computing and

Emerging Technologies. ICCET 2023. Communications in Computer and Information Science, vol 2056. Springer, Cham. https://doi.org/10.1007/978-3-031-77620-5_20

Meeran, M. T., Raza, A., & Din, M. (2018). Advancement in GSM Network to Access Cloud Services. Pakistan Journal of Engineering, Technology & Science [ISSN: 2224-2333], 7(1).

Raza, Shoukat, N., Salahuddin, Aslam, M., & Shahzad, I. (2024). DETECTION OF DIABETES APPLYING MACHINE LEARNING TECHNIQUE. Kashf Journal of Multidisciplinary Research, 1(11), 107-124

Mahmood, F., Abbas, K., Raza, A., Khan,M.A., & Khan, P.W. (2019). Three Dimensional Agricultural Land Modeling using Unmanned Aerial System (UAS). International Journal of Advanced Computer Science and Applications (IJACSA) [p-ISSN : 2158-107X, e-ISSN : 2156-5570], 10(1).

Aslam, M., Salahuddin, Ali, G., & Batool, S. (2024). ASSESSING THE EFFECTS OF BIG DATA ANALYTICS AND AI ON TALENT ACQUISITION AND RETENTION. Kashf Journal of Multidisciplinary Research, 1(11), 73-84

Syed Shahid Abbas, Salahuddin, Abdul Manan Razzaq, Mubashar Hussain, Meiraj Aslam, Prince Hamza Shafique, & Muhammad Asif Nadeem. (2024). Optimized AI-Driven Intrusion Detection in WSNs: A Semi-Supervised Learning Paradigm. Journal of Computing & Biomedical Informatics.

Hekmat, A., Zuping, Z., Bilal, O., & Khan, S. U. R. (2025). Differential evolution-driven optimized ensemble network for brain tumor detection. International Journal of Machine Learning and Cybernetics, 1-26.

Salahuddin, Hussain, M., hamza Shafique, P., & Abbas, S. S. (2024). INTELLIGENT MELANOMA DETECTION BASED ON PIGMENT NETWORK. Kashf Journal of Multidisciplinary Research, 1(10), 1-14.

Khan, S. U. R. (2025). Multi-level feature fusion network for kidney disease detection. Computers in Biology and Medicine, 191, 110214.

Khan, S. U. R., Asif, S., & Bilal, O. (2025). Ensemble Architecture of Vision Transformer and CNNs for Breast Cancer Tumor Detection From Mammograms. International Journal of Imaging Systems and Technology, 35(3), e70090.

Syed Shahid Abbas, Salahuddin, Abdul Manan Razzaq, Mubashar Hussain, Meiraj Aslam, Prince Hamza Shafique, & Muhammad Asif Nadeem. (2024). Optimized AI-Driven Intrusion Detection in WSNs: A Semi-Supervised Learning Paradigm. Journal of Computing & Biomedical Informatics.

Khan, S. U. R., & Khan, Z. (2025). Detection of Abnormal Cardiac Rhythms Using Feature Fusion Technique with Heart Sound Spectrograms. Journal of Bionic Engineering, 1-20.

Salahuddin, Syed Shahid Abbas, Prince Hamza Shafique, Abdul Manan Razzaq, & Mohsin Ikhlaq. (2024). Enhancing Reliability and Sustainability of Green Communication in Next-Generation Wireless Systems through Energy Harvesting. Journal of Computing & Biomedical Informatics.

Khan, M.A., Khan, S.U.R. & Lin, D. Shortening surgical time in high myopia treatment: a randomized controlled trial comparing non-OVD and OVD techniques in ICL implantation. BMC Ophthalmol 25, 303 (2025). https://doi.org/10.1186/s12886-025-04135-3.