International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Implementing Spectral-Domain Feature Mapping for Sketch-Based Image Retrieval

K. Durga Prasad[1*], K. Manjunathachari[2], M. N. Giri Prasad[3]

Submitted: 02/12/2023 **Revised**: 29/01/2024 **Accepted**: 04/02/2024

Abstract: Sketch-based image retrieval (SBIR) has grown significantly in popularity across various real-time applications that use automated image processing. Examples of applications that use SBIR include banking, internet searching, and secure coding. On the other hand, these applications have a higher need for performance in terms of speed and accuracy. Because of the high cost of testing and the limited resources, the representative traits have to be narrowed down to be more selected for descriptive purposes and lowered in the count. The existing methods have low computational speed, more complexity and less accuracy. This paper presents a novel spectral deviation method in a sketch-based picture retrieval method. Spectral Coding Selective Feature Mapping (SpecCode SFM) is the name of the suggested technique. It was created by comparing the spatial correspondence between sketch images and their raw image counterparts. The suggested approach for retrieval passes the image as a free-hand sketch processing. This method tolerates scale and orientation and provides a sizable retrieval efficiency. The developed approach's findings are compared to those of cutting-edge techniques, and it is discovered that SpecCode SFM achieves 99.87% accuracy, 0.87 detection rate, 0.89 MCC, and 0.12 sec of computing time.

Keywords: *Sketch-based image retrieval; spectral map coding; true and false regression; Gabor filter; content distortion*

1. Introduction

Digital media is becoming easier to create and more popular. The availability of digital cameras, scanners and other devices to digitize what we see is increasing. This, combined with the decreasing cost and rising availability of digital storage mediums, has led to vast growth in the size and commonality of image and video collections [1]. Large data collections lead to problems with organization and accessibility, resulting in the need for suitable search capabilities. Searching for digital media is a problem that has existed for a while and has yet to be universally accepted. Systems that allow digital video and photo library search commonly rely on meta-data for retrieval [2]. Sites like Google Image Search or YouTube search their databases using captions, comments, and surrounding text. Reliance on the existence and accuracy of this data is where these systems are flawed. This data needs to be entered and entered accurately for the results of a search to be useful. Content-based image and video retrieval focus on retrieving images based on the visual content of the images themselves [3]. Image processing has been used in different areas of applications, such as military applications, navigation, satellite image processing, medical application, forensic science, the banking sector, security system, etc. The applications greatly benefit from the advancement of rapid development in image processing. With so much available visual information, image retrieval becomes increasingly crucial. Retrieving a photo from a collection using one or even more evaluation metrics is known as image retrieval [4]. The majority of existing systems—Google Online Search, becoming a good illustration that retrieval utilizing information from the meta-data of the collection's pictures. Numerous photographs are available without the accompanying text, comments, or descriptions that make up this meta-data. The compilation of this metadata will be an extremely time-consuming and costly task for huge image collections.

Numerous techniques for Content-based image retrieval (CBIR) have been tried and tested, but as noted in several studies, there needs to be a widely recognized strategy [5]. Such techniques concentrate on parameterizing and extracting the visual information in the image. The main focus of these techniques is the recovery of colour, structure, geometry, or a mixture of the three. Using the comparison methodologies, they chose and the domains of information retrieval on which they were tested, each of these is demonstrated by their various publications to generate remarkably reliable data. However, the areas of the test were not all consistent, so the results produced by many of the papers cannot be compared directly. One conclusion that can be drawn from this is that different techniques produce better

[1]*Department of ECE, JNTUA, Ananthapuram, AP, 515002, India.*
Email: kalasapatidurga@gmail.com
[2]*Department of ECE, GITAM University, Hyderabad, Telangana, 502329, India*
Email: manjunath4005@gmail.com
[3]*Department of ECE, JNTUA, Ananthapuram, AP, 515002, India*
Email: mahendragiri1960@gmail.com

results depending on the type of retrieval and search. A query by sketch system can function to a higher degree of accuracy using a different technique to produce the highest accuracy with a query by example image system [6].

Images are characterized by the outrebound edge area of the picture when it comes to defining sketch-based image retrieval. This characteristic determines the context of the picture retrieval and the bounding boundaries of the image in a rough sense. The image's form is defined by the edge features, which also represent the curvilinear characteristics of the test sample. The user is responsible for supplying the drawing for the sketched-based retrieval method to locate a picture. Users can get visual information on a picture that they are looking for. However, in order for the system to carry out a successful recovery, it must supply essential visual information [7]. The information in a picture can be broken down into three categories: colour, texture, and geometry. Outlines in a sketch convey shape, while parts of a single colour provide vital information on roughness and colour arrangement. The drawing's visual components crudely represent these crucial details.

The strokes' direction and the lines' uniformity in the sketching could be leveraged by an optimized feature descriptor in the SBIR domain to obtain contextual characteristics from hand-drawn images and function adequately on realistic photos. Additionally, a component description must be robust enough to handle various kinds of image data and resolve sketching ambiguity brought on by differences in people's drawing styles and preferences. Applications such as banking, internet searching, and secure coding have a greater need for performance in terms of both speed and accuracy. This is because these applications deal with sensitive data. Because of the high cost of testing and the restricted resources, the list of representative characteristics has to be whittled down to be more selective for descriptive purposes and reduce the total number of characteristics. The currently available procedures have a slow computing speed, a higher level of complexity, and a lower level of accuracy.

An appropriate appraisal criterion should thus be capable of clearly and quickly evaluating any such variance. The motivation of SpecCode spectral feature mapping (SFM) is to use the system with limited resources to get more accuracy and high computational speed which the following contributions can achieve. The contributions of this work are as follows:

- Initially, a Gabor-based Information Displacement Assessment is offered to reduce the background noise in characteristics. This assessment constantly superimposes the boundary screening findings of Gabor blur processing.
- The proposed method uses the image's outside and inner contexts to pick out certain portions as bounding boxes.
- The feature selection procedure involves matching the context of mapping the edge area of a drawing with a realistic picture and then using regression to reduce the difference between the two images' domains.
- The regression iteration is chosen based on some surrounding images with higher energy spectral values.
- The proposed method demonstrated improved accuracy and reduced feature dimension compared to prior spectral codings methods like HoG and other deep learning methods.

This research article is structured as follows: In this introductory portion 1, a thorough introduction to digital technology and image analysis is covered. Section 2 discusses the literature overview of earlier work on Sketch-based picture retrieval. Section 3 provides background about Sketch-Based Image Retrieval. Section 4 presents the suggested technique for effective image retrieval with a conceptual diagram. Section 4 of the proposal discusses the project's findings. Finally, the new work in section 5 brings this study piece to a close.

2. Related Works

The author in [8] developed SceneSketcher-v2, a Graph Convolutional Network (GCN)-based design, which bridges the category barrier using a quadruplet training regime, end-to-end learning, and a meticulously crafted graph convolution network to combine the multi-modality data in the query sketches with accurate shots. Within [9] provided a new deep SBIR framework that attempts to estimate concentration mappings across temporal and spatial dimensions, improve channel attention and spatial attention modules, and investigate Methods to improve the model's capacity for creating and comprehending spatial sequence information. Within [10], a paradigm for sketch production should be first put forth to take the role of the traditional preprocessing step of approximately separating picture edges. In addition, this concept could solve the problem of sketching information shortage. Then, using a deformable convolutional neural network and simultaneously taking into account semantic features, we build a unique fine-grained sketching image-based retrieval (FG-SBIR) approach. To simultaneously enhance sketch-to-photo synthesis and image retrieval, [11] proposes an Approaching-and-Centralizing Network (termed

"ACNet"). In order to acquire domain-agnostic interpretations and classification fairly obvious for generalizing to unknown subcategories, the fetching modules direct the synthesizing generator to produce huge quantities of varied picture-like images that progressively approximate the photo area.

To solve this problem, [12] suggests combining methods based on deep learning and image preprocessing. A binary picture highlighting the edges in the raw image is produced using the Canny-Edge detection algorithm. A Convolutional Neural Networks (CNN) model built on ImageNet extracted the in-depth features. Cosine similarity and Euclidean distance measurements are used to create the rank list of potential natural image candidates. In [13]'s innovative cross-modal retrieval problem of fine-grained instance-level sketch-based video retrieval (FG-SBVR), a sketch sequence is utilized as a request to recover a particular target video instance. This is more difficult than sketch-based still information retrieval or coarse-grained category-level video retrieval since visual appearance and motion must be matched simultaneously. They provide the initial richly annotated FG-SBVR dataset. To choose the optimal perspective for a shape, [14] proposes a learning framework with contour recovery utilizing two Siamese VGG-16 Convolutional Neural Networks (CNNs) and a Convolution neural composite prediction model. In this approach, categorization functions are carried out by the AlexNet architecture, while the VGG-16 Convolution layer handles background subtraction.

Additionally, the resemblance between the output features from the two Siamese networks is evaluated using a features segmentation technique. A graph CNN-based component that spreads the conceptual category hierarchy to the communal space is described in [15]. They continue to investigate the potential of encoding the facility in elements of hash codes to provide a faster reaction time throughout the interpretation.

Re-Ranking Algorithm That Improves Sketch-Predicated Image Retrieval Performance in [16] (SBIR). Again, from existing approaches, the suggested scheme may strengthen homogeneity feature assessment between images using category data supplied by CNN. One CNN model learns for the demotion of drawings, while another is prepared to relegate raw images to accomplish efficient demotion. Deep learning can capture the semantics of both frames by building dual CNN models. [17] proposes an SBIR re-ranking strategy that utilizes multi-clustering. We used blind evaluation to create the re-ranking strategy in an automatic manner that is adaptable to many kinds of image datasets and undetectable to users. In contrast to existing methods, this re-ranking methodology utilizes the semantic information of 3 distinct types of images: edge maps, item images (natural photographs with salient items and a black backdrop), and instinctual photographs. In [18], Text retrieval and image retrieval are the two main steps of the proposed natural language processing with deep learning enabled hybrid content retrieval (NLPDL-HCR) and image retrieval (IR). The suggested NLPDL-HCR model incorporates an ideal gated recurrent unit (GRU) model with a term frequency-inverse document frequency vectorizer throughout the text retrieval process. In [19], put forth an instant design that begins retrieving when the user draws. To do this, they create a cross-modal retrieval framework based on reinforcement learning that optimises the ground-truth photo's rank throughout an entire sketch drawing session. The findings above confirmed that it was challenging to separate the attributes that belong to distinct boundaries from the patch's overall feature space. The efficacy of current approaches for SBIR suffers across scale, translational, or rotations variations, and evidenced image retrieval systems have two drawbacks, including the need for a significant amount of manual effort when the features of digital are vast. Human annotators struggle to accurately describe an image's rich content because of perceptual imperfection. Table 1 shows the literature review summary.

Table 1. Literature review summary

Author	**Method**	**Benefit**	**Issue**
Liu et al. (2022)	SceneSketcher-v2	Easy & effective	Extremely scale-sensitive, requiring a low-level treatment
Chen et al. (2022)	Deep SBIR model	Lower dimensional depiction in a domain.	Estimating unsupervised parameters is challenging.
Zhang et al. (2022)	Fine-grained sketch-based image retrieval (FG-SBIR)	It is simple to identify limitations	We need more training case studies.

Ren et al. (2021)	Approaching-and-Centralizing Network	It is simple to evaluate the connection across face sketches and facial pictures.	At most, one Sketch can be downloaded.
Kumar et al. (2021)	ImageNet based CNN	Improve the original image's quality	Low-level characteristics need help to represent and meaningfully comprehend the visual background.
Yu et al. (2021)	Fine-grained instance-level sketch-based video retrieval (FG-SBVR)	The good recognition rate for sketch photos	Low-level characteristics cannot represent and meaningfully comprehend the visual environment.
Zhou et al. (2019)	Siamese VGG-16	The best method for efficiently categorizing and finding many photographs	Unusable for regular users because they must supply their search terms as photos
Chaudhuri et al. (2020)	Graph CNN	Eliminate uncertainty	Time-consuming, costly, and arbitrary
Anisha et al. (2021)	Dual CNN Models	Taking fewer moment	Practically useless for tagging massive image databases
Wang et al. (2019)	Re-ranking approach	Greater consistency	Greater time commitment

3. Proposed Methodology

3.1 Background of Sketch-Based Image Retrieval

The example image is often an image with content matching the image's content that the user wishes to retrieve. However, it has also been proposed that the users of a content-based image retrieval system could search by entering a sketch of the image or content of the image or images they wish to locate. Charles E. Jacobs et al. present one such system in their paper on Fast Multi-resolution image Querying**.** This form of Querying is known as Sketched-Based Content-Based Image Retrieval. Querying by image content (QBIC) is one such system with functionality for sketched-based retrieval [20-25]. Sketched-based content-based image retrieval offers several advantages for searching large image databases. Many systems require the user to search by an example image. This image needs to be similar to the image the user wishes to locate. Finding such a visually similar image can often become time-consuming and, in the worst case, reduce the task of finding the desired image down to the laborious task of browsing the database. Furthermore, should database access be across the internet, slow or busy networks could make browsing the database very time-consuming. In this case, a sketched-based system would allow users to input a sketch representing the images they wish to retrieve [26-32].

Several systems allow the user to enter visual information using various metrics, such as the percentage of a certain colour or contrast information. Sketched input allows a user a much more intuitive method of entering such data, including much more information about the layout and spatial information of the colour and shapes within the image. Providing the system with more accurate and detailed information allows the sketch-based retrieval system to produce more accurate and reliable results. Many techniques have been suggested to make this form of retrieval possible. These techniques require extracting visual information from the database images and the sketched image for comparison. Within an image, the human visual system can detect a vast amount of data, which must be categorized and represented compactly [33-36]. The data stored within an image can be separated into colour, texture, and shape information, all of which have several possible representations, all of which can be used to compare sketched images to images within a database. One of the key considerations must therefore be the environment in which the technique is applied. For query by Sketch, it is important to consider the properties that a sketched image possesses when compared to a photograph. The level of detail within the Sketch will be much lower than that of a photograph or scanned image as used in a

query-by-example style of system. Sketched input is likely to represent areas of texture as areas of consistent colour, colours are likely to vary dependent upon the user's memory, and visual interpretation and shape are subject to the affine transformation from sketching error. Arguments exist for and against retrieving each type of data from the image [37-40].

Colour-based techniques provide a much lower level of implementation and computational complexity. However, they are likely to produce less accurate results if we consider the properties of the sketched input. The sketched input is drawn from the mind of the human. So, the memory of colours could often be warped to the extent where very different colours are present in the Sketch compared to that of the target image or images. Colour data would therefore prove inaccurate for comparison purposes, particularly with sketches. Queries produced from memory. The shape of the areas of the colour may be remembered much more clearly, even if the area's colour is completely different from that intended. However, this colour difference would only be relevant if shape extraction was used for comparison since only the shape of the areas is considered and not their colour. Shape extraction presents a much higher level of implementation and computational complexity due to the nature of the techniques employed. Shapes would, however, be subject to affine transformation resulting from errors within memory and with the production of Sketch. Depending on the user's artistic skill, these errors will vary. Therefore, any shape extraction technique should also allow comparison independent of the affine transformation [41-46].

In [47] provide a Gaussian blur-based multiscale edge extraction (GBME) approach to continually superimpose the edge filtering results after Gaussian blur processing to capture more specific and detailed features. A hybrid barycentric feature descriptor Randomly Sampled with Barycenter Histogram of Oriented Gradient (RSB-HOG) that extracts HOG features by randomly selecting points on a sketch's edge is the second method we developed. Additionally, we enhance the feature descriptor's representational power by incorporating the directional distribution of all sampling points' barycenters, which will better capture the semantic content of contours. In [48], a deep, fully convolutional neural network is a general and personalized model using transfer learning to achieve fine-grained image semantic features.

Texture-based techniques could need more detail in a sketched image since texture revolves around recognizing areas of the repeated pattern. Areas of consistent colour will often represent such areas within a sketch. However, texture-based techniques represent areas of consistent colour very well, and such techniques should be used for sketch-based content-based image retrieval. Texture-based techniques also offer high degrees of complexity compared to colour, but the implementation is generally less complex than shape. An edge-guided regression model is presented in developing a sketch-based image retrieval. This approach defines an edge-based cross-domain approach lowering the feature description gaps. The approach defines a shape-based regression for feature description, which defines a similarity property of Sketch and natural images.

3.2 Proposed Sketch-Based Image Retrieval System

In this approach, a mapping method based on feature mapping is developed based on feature space. This approach developed a retrieval performance based on three branches, defended as "sketch branch, positive branch, and negative branch."

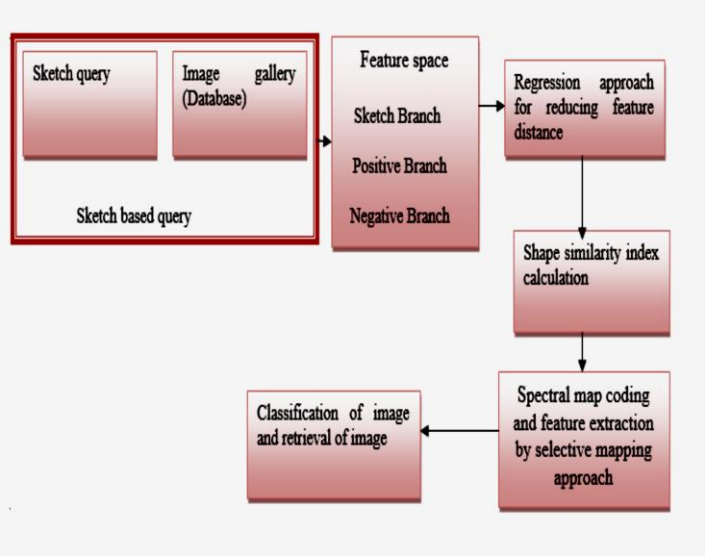

Fig 1. Block diagram representation of Sketch-based image retrieval system

The images with a synonymous label and defined feature are passed to a positive link. A negative link of the SBIR system defines the images which differ in feature from the label. The sketch branch is an independent branch where the two links are interlinked to formulate this branch. The regression approach minimizes the feature distance by deriving the shape similarity between the sketch and raw image representations. The regression is developed to minimize the feature gap by correlating shape information's negative and positive branches, as shown in Fig.1.

3.3 Analysis of Regression Loss

The spectral resolution is independent in developing the feature, and the correlative band results in large redundant features. This redundant feature increases the overhead of the system. This affects the effectiveness of feature representation for the edge map of the test image. In minimizing the regression loss, the edge descriptor is defined in an 8x8 dimension, which is interpolated to a linear scale as 64 value vectors. The loss of regression in the proposed approach is defined by,

$$L_{sr} = \left| f_{ShReg}^{Q}(Q), m^{Q}edge \right|^{2} \quad (1)$$

Where,

L_{sr} is defined as the loss of regression model.

f_{ShReg}^{Q} is the shape regression of the SBIR unit

$m^{Q}edge$ is the targeted shape vector

The regression loss makes the edge details retain and minimizes the information loss in defining the features. The regression has an overall loss by cross-domain processing given by,

$$L_r(Sk, Ps, Ng) = \mu \sum_{Q1=Sk,Ps,Ng} Lc(Q) + \beta Lt(Sk, Ps, Ng) + \gamma \sum_{Q2=Ps,Ng} L_{sr}(f_{ShReg}^{Q}(Q2)) \quad (2)$$

Here, $Sk, Ps, and\ Ng$ are the sketches, positive and negative edge features. $Lc(Q)$ is the loss during the classification process. L_{sr} defines the loss during the regression process, and the losses in the link process are given by Lt. The loss of regression is considerable in the case of a redundant spatial feature. A spectral map-based feature representation is proposed to overcome the processed feature overhead.

3.4 Proposed Spectral Map Coding

The variation of the line's alignment in a preliminary test representing the tissue section of a corresponding point is referred to as substance deformation. Horizontal deviations are, therefore, the most interesting discovery in the system to detect alteration of content. When distortion analysis is processed using various processing techniques, interference plays a vital part in the individual's correctness. The background distortions reduce the detection performance of information deformation in most processing approaches. The intrinsic ability of the Gabor filter to define textural features in line alignment despite various magnification and orientation effects. It is suggested to apply a Gabor filter for material deformation in which the linear component's directing breakdown is applied in a multi-resolution region to recover the necessary topographical data. Using a 2D-labour filter provided by the Gabor-based Contents Displacement Assessment gets processed.

$$\gamma(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{{\sigma_x}^2} + \frac{y^2}{{\sigma_y}^2} \right) + 2\pi jWx \right] \quad (3)$$

The Gabor kernel, which represents the sinusoid's amplitude and thus the Stochastic envelope's sample variance, is a complicated sinewave manipulated by Gauss. The photograph is handled during the filtering stage to have the DC responsiveness be zero and specified for a minimally duplicated depiction. This is accomplished by rotating and dilating the (x,y) and generated by the output function.

$$\gamma(x', y') = [-(x - x_0)sin\theta + (y - y_0)\cos\theta] \quad (4)$$

Here, x_0 and y_0 are located in the filter's middle. The inclination is specified for the direction as n/K, wherein n is the necessary inclination, and K is the scale parameter. The Gabor filter's power spectrum equivalent is provided by

$$\vartheta(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\frac{(x-W)^2}{{\sigma_x}^2} + \frac{y^2}{{\sigma_y}^2} \right) \right] \quad (5)$$

When using the Gabor filtration system, the machine detects when the Gabor results deviate in a particular way and declares an alignment choice as Relevant interpretation. However, this method maps the direction to find the alternative, which could be interpreted as noise-induced distortion. Therefore, in this circumstance, this detection results in misclassification. The projection is produced in the spectrum domain using energetic mappings in conjunction with advanced to prevent misdiagnosis.

3.5 Spectral Deviation Mapping

Every direction of a Gabor filter is treated with an energetic calculation determined by the spectrogram of every direction group to identify material deformation based on signal projection. Every scaling tier's direction is determined by,

$$\theta(x, y) = -\frac{\pi}{2} + \frac{\pi}{K} \quad (6)$$

Screening is used to handle the elements of the direction for spectrum estimation, and a Gaussian is built to screen the radiation element described by,

$$h(k) = \frac{1}{2}\left(h_0\left(\frac{\theta}{2}\right) + h_0\left(\frac{\theta-2\pi}{2}\right)\right) \quad (7)$$

Every alignment band's spectroscopic element is derived using harmonic filtering. Diagram Fig. 2 illustrates the signal bandwidth for a corresponding point.

The testing region's spectrum depiction shows how the spectroscopic component varies in the area under consideration; this difference in the function of independent variables can be seen based on the brightness of the test image's pixel's location.

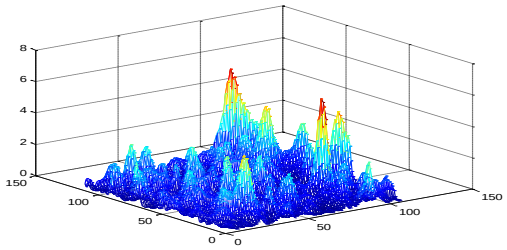

Fig 2. Spectral plot for a test image

Within every scenario, the image is transformed according to predetermined rotations while the spectrum intensity is calculated and stored. Each region is analyzed for spectral mapping in the projection procedure, and an area with a significant spectroscopic divergence is delayed. The divergence among many orientations is derived from the spectrum variation. The band differential with the greatest deviation for each relationship is considered substance deformation. Following segmenting, the observed Content distortion region is analyzed and sent to the decoding unit for noise or Content deformation classification. A sample picture's spectrum graph is depicted in Fig.2.

3.6 Selective Feature Mapping Approach

The mapping method uses Selective Feature Mapping for spatial property reduction, in which the directions of projections are obtained and exploits the full scatter in every picture category. This projection needs to be revised for the picture's reconstruction based on its dimension. In this methodology, the between-class scatter is not used. The projection is optimal from various categories with discrimination. Consider ST is the full scatter matrix outlined in equation (8)

$$S_T = \sum_{k=1}^{N}\left(T_k - \psi\right)\left(T_k - \psi\right)^T \quad (8)$$

The projection $W_{opt}$is selected for the maximization of the determinant of the projection sample's total scatter matrix as shown in equations (9) and (10)

$$W_{opt} = \arg\max_{W}\left|W^T S_T W\right| \quad (9)$$

$r = [w_1, w_2, \ldots\ldots, wi]$ (10)

where a set of *n*-dimensional vectors of S_T *is represented by* $\{w_i|\ i=1,2\ldots\ldots,m\}$ concerning the *m* largest eigenvalues. In the database, sketch classes are represented by X1, X2, …, and Xc. Xi. i = 1, 2, …, c is the vector representing every Sketch with k point images Xj, j=1, 2, …, k. The mean image μi of every class Xi as shown in equations (11) and (12)

$$\mu(i) = 1/k \sum_{j=1}^{k} \binom{n}{k} x\ (j) \quad (11)$$

$$S\ (b) = a_0 + \sum_{i=1}^{c} N\ (i)(\mu\ (i) - \mu\ )\ (\mu\ (i) - \mu)T \quad (12)$$

A generalized problem of eigenvalue wherever the Z's columns are provided by the vectors Zi as shown in equation (13)

$$S\ (B)\ Z(i) = \lambda(i)\ S(W)Z(i) \quad (13)$$

The multiplication of $S(W)^{-1}$ and S(B) are used to calculate eigenvectors, and then the feature space dimension is reduced. At last, matrix Z is composed of $S(W)^{-1}$ * S(B) eigenvectors are represented by introducing every eigenvector Z(i) present in the Z column

Training:

Initiate total samples (N)

For N={1,2,…n},

Pre_N=pre{1,2,…n}

Read n image

Resize n= $n_{256\times256}$<=k

Initiate Feature selection;

Decompose k

Compute PSD(k)<=L

L= Th = max(C)/2,

where max(c) is the maximum coefficient value of a band

Sel=coef(k), when> Th

Iterate i=1,2,..n and save as ∅

Database formation;

Create feature vector,

Buffer selected feature mapped with the image index

Create an image array mapping the image to its corresponding index

End

Testing:

Read query sketch

Preprocess for the uniform dimension of 256x256

Extract features by selective mapping approach

Classifier;

Compute distance vector (D) by mapping of database feature with the query feature

Output classified image with min(D).

4. Results & Discussions

Dataset Description- Using a free hand-drawn sketch image prepared for retrieval utilizing the SBIR system, the evaluation of the suggested approach is validated. An assortment of learning images from the coil-100 database are used in the examination. This database is a collection of photos of frequently used objects that were taken from various angles. Each database image is treated for form detection as part of the developed system's learning process, after which the image is processed for edge detection to produce a closed contour.

Implementation details- In the process of image coding, each of the images is processed for edge detection, where the image is processed for spectral decomposition and selecting edge regions are processed as by the proposed approach. The test is made based on a random selection of 10 test samples sketched as input, and the system is developed for the top 4 classified outputs. The suggested method was created using a MATLAB tool, which reads and processes the test instance for a consistent 256x256 image dimension. Filtering for -/2 to /2 Scale alignment was done using 12 treatment rounds.

The testing is developed, and the feature dimensions of a selected feature defined by the number of selected feature values are computed. The proposed approach is compared to the existing state-of-art methods, namely, Histograms of Oriented Gradients (HOG) [18], Gradient Field (GF)-HOG [19], Soft Computation of the Histogram of Edge Local Orientations (SHELO) [20] and Learned

KeyShapes (LKS) [21] methods. Wherein deep learning approaches such as Siamese- AlexNet [22], Triplet-AlexNet [22], Deep Sketch Hashing (DSH) [22], 3D shape [23], SaN [24], Siamese Convolutional Neural Network (CNN) [25] and GN Triplet [26] are also evaluated to validate the performance by comparing with Proposed Method (SpecCode_SFM). The evaluation is made based on Mean Average Precision (MAP) as the retrieving performance given by,

$$MAP = \frac{\sum_{n=1}^{N} Avg(\mathrm{Prc}(n))}{N} \quad (14)$$

Where *Prc*(*n*) defines the retrieval precision, and n is the test sample. N defines the number of test samples in learning, and *Avg(:)* function is the averaging function.

Table 2. Observation of MAP variable for the developed approach

Method	Dimension	MAP
HOG [18]	1296	0.091
GF-HOG [19]	3500	0.119
SHELO [20]	1296	0.123
LKS [21]	1350	0.157
Siamese Alex Net [22]	4096	0.367
Triple Alex Net [22]	4096	0.448
DSH [22]	128	0.570
3D shape [23]	64	0.072
SaN [24]	521	0.154
Siamese CNN [25]	64	0.322
GN Triplet [26]	1024	0.187

Proposed Method	56	0.743

Table 2 shows the observation of the MAP variable for the developed approach. The observations in Fig.3 illustrate a feature dimension count of 56 for the proposed approach. Wherein the HoG approach has 1296 selected features, the proposed approach has 1240 features.

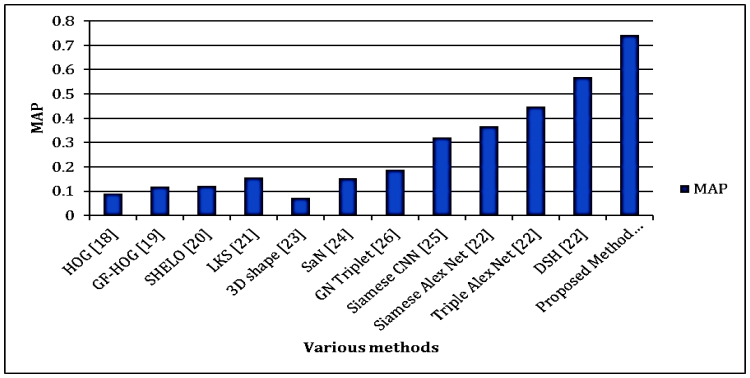

Fig 3. Analysis of Mean Average Precision

The particular features are non-redundant and descriptive, which results in the procession of 0.743, which is comparatively higher than the existing approach. The distinct selective features increased precision, minimizing the classification error of the proposed system. The simulation result over the test inputs is shown in the figure below. The test image for the processing is illustrated in Fig.4 below.

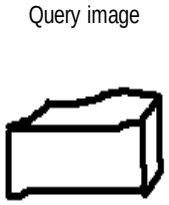

Fig 4. Test sample

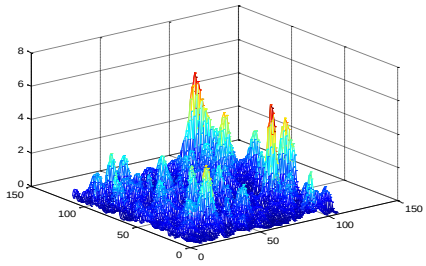

Fig 5. Spectral plot for the reference model

To obtain sharp features for the speculation in the Test sample, the image is normalized for intensity utilizing normalization, and the intensity is normalized for low intensities. The picture is then treated for computing content deformation. Block analysis creates a reference architecture and a test model from the image. Depending on the area of the sample set with homogeneous frequency response, a standard model is created. A power spectral density computation is utilized to calculate the spectrum complexity for the model, measuring the spectral density over the recovered area with the PSD(x) function. Fig.5 below displays the standard model's spectral pattern.

A spectrum constituent is calculated for every unit in the central section, and a spectrum graph for a corresponding point is shown in Fig.7, given below. A sample query passed with an orientation of 90^0; the result is shown in Fig.6.

Fig 6. Query sample at 90^0 orientation

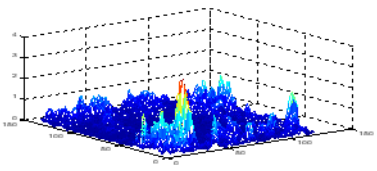

Fig 7. Spectral plot for a test region

The area of concern with Content deformation would then be found using spectrum projection. Fig.8 below provides a spectrum mapping example.

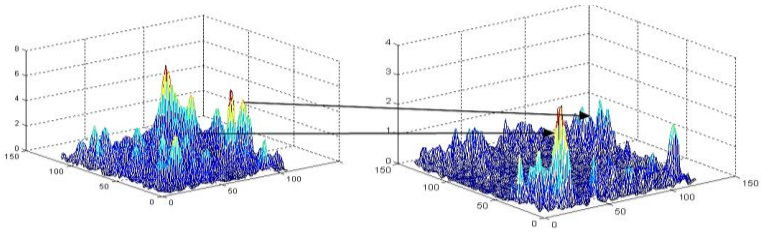

Fig 8. Spectral correlation approach

Classified samples for the given sample query are shown below in Fig.9.

Fig 9. The top 4 classified observations were obtained from the database for the query sample at 90 ^{0} orientation.

Reliability, detection rate (DR), Matthew's correlation coefficient (MCC), and computational burden are used to calculate the decision performance of the suggested approach. Mentioned Table 3 below lists the assessment of the suggested technique.

Table 3. Comparison of detection performance for the developed approach

Samples	Method	Accuracy (%)	DR	Mcc	CT(s)
S1	Spectral mapping	99.85	0.72	0.72	0.347383
	Regression model [27]	99.22	0.58	0.58	0.138035
S2	Spectral mapping	99.52	0.60	0.60	0.259342
	Regression model [27]	99.13	0.88	0.87	0.147938
S3	Spectral mapping	99.76	0.72	0.73	0.358719
	Regression model [27]	99.25	0.90	0.85	0.147317
S4	Spectral mapping	99.74	0.72	0.72	0.366849
	Regression model [27]	99.56	0.90	0.87	0.178043

Fig.10 depicts the accuracy comparison of an existing regression model with the proposed spectral mapping. The X and Y axes show the number of samples and the accuracy values obtained in percentage, respectively. When compared, existing regression model methods achieve 99.10% of accuracy, respectively, while the proposed spectral mapping method achieves 99.87 % of accuracy, which is 0.77% better than the regression model.

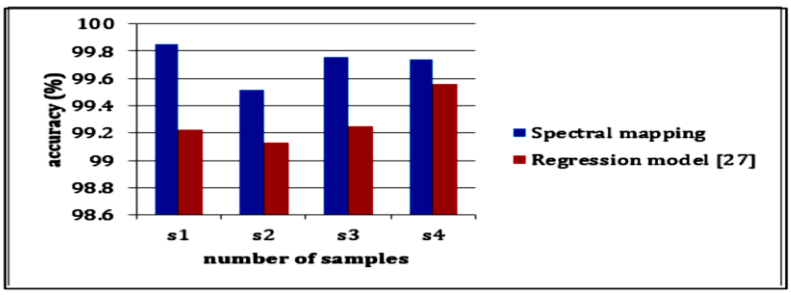

Fig 10. Comparison of accuracy

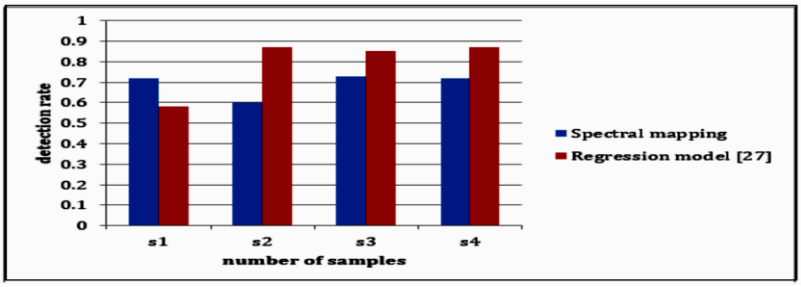

Fig 11. Comparison of detection rate

Fig.11 depicts the detection rate comparison of an existing regression model with the proposed spectral mapping. The X and Y axes show the number of samples and the detection rate, respectively. When compared, existing regression model methods achieve 0.60 respectively, while the proposed spectral mapping method achieves 0.87 respectively. Fig.12 depicts Matthew's correlation coefficient comparison of an existing regression model with the proposed spectral mapping. The X and Y axes show the number of samples and the correlation coefficient, respectively. When compared, existing regression model methods achieve 0.57 respectively, while the proposed spectral mapping method achieves 0.89 respectively.

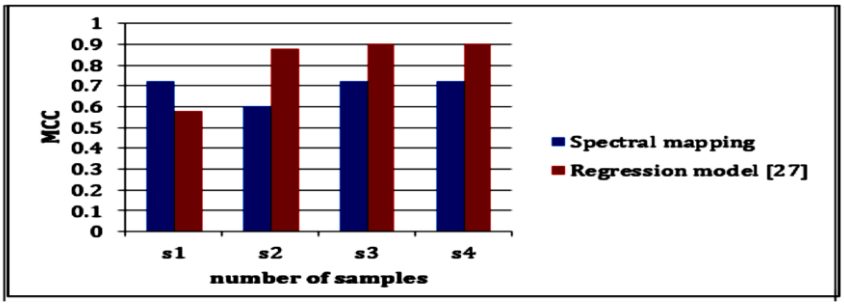

Fig 12. Comparison of Matthew's correlation coefficient

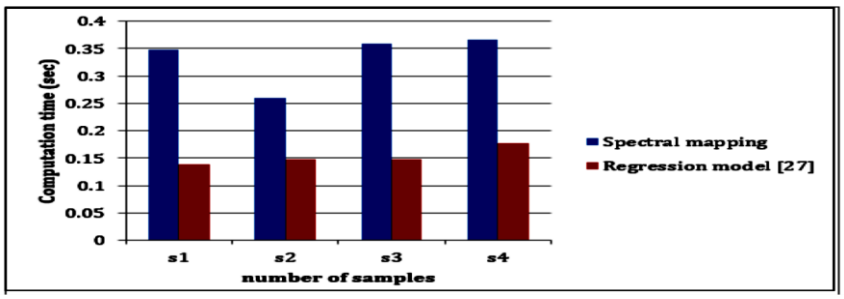

Fig 13. Comparison of computation time

Fig.13 depicts the computation time comparison of an existing regression model with the proposed spectral mapping. The X and Y axes show the number of samples and the computation time, respectively. When compared, the proposed model achieves less time.

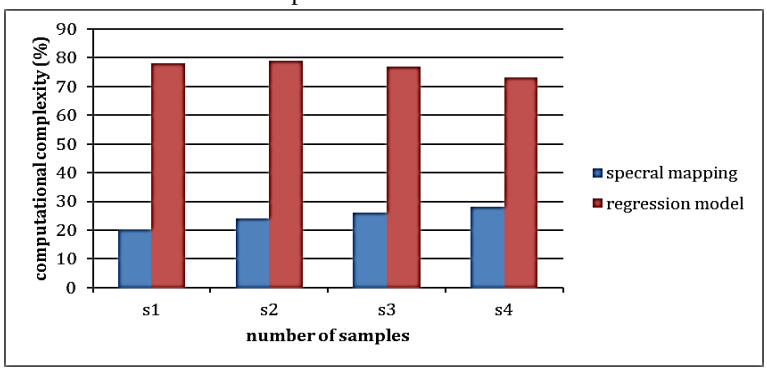

Fig 14. Comparison of computational complexity

Fig.14 depicts the computational complexity comparison of an existing regression model with the proposed spectral mapping. The X and Y axes show the number of samples and the computational complexity, respectively.

When compared, the proposed model achieves less complexity.

The main drawback of the CBIR system is that the images with similar low-level features may differ from the query picture in positions of the semantics perceived by the user

5. Conclusion

The dimensional overhead of feature selection in image retrieval is addressed in this paper. The proposed method is developed as a new spectral selective method, where the features are selected based on the energy dominance of the image region. This approach selectively selects the bounding regions of images at a given image's outer and inner contexts. In the feature selection process, the context of mapping the edge region of a sketch is matched with a natural image, and regression is developed to minimize the cross-domain gap between the two images. The regression iteration is selected for selective higher energy spectral coefficients rather than all bounding pixels. The approach illustrated a higher precision and lower feature dimension than existing spectral coding approaches such as HoG and various deep learning approaches. Future work concentrates on including deep learning methods with metaheuristic feature selection.

Acknowledgments: The authors would like to thank the R&D departments of JNTUA, Ananthapuram, AP, 515002, India, for supporting this work.

Funding Statement: The authors received no specific funding for this study.

Author Contributions: study conception and design: K.D. Prasad; data collection: K.D. Prasad; analysis and interpretation of results: K.D. Prasad, K.M. Chari. M.N.G. Prasad; draft manuscript preparation: K.D. Prasad. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: The data and materials are available from the corresponding author upon request.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] M. Eitz, K. Hildebrand, T. Boubekeur and M. Alexa, “Sketch-based image retrieval: benchmark and bag-of-features descriptors,” *IEEE Transactions on Visualization and Computer Graphics*, vol.17, no.11, pp.1624-1636,2010.

[2] R. Hu and J. Collomosse, “A performance evaluation of gradient field hog descriptor for sketch-based image retrieval,” *Computer Vision and Image Understanding*, vol.117, no.7, pp.790-806,2013.

[3] J.M. Saavedra, J.M. Barrios and S. Orand, “Sketch based image retrieval using learned keyshapes (lks),” *British Machine Vision Association and Society for Pattern Recognition*, vol. 1, no. 2, pp. 7,2015.

[4] L.Wang, X. Qian, Y. Zhang, J. Shen and X. Cao, “Enhancing sketch-based image retrieval by cnn semantic re-ranking,” *IEEE Transactions on Cybernetics*, vol.50, no.7, pp.3330-3342,2019.

[5] Y. Li and W.Li,“A survey of sketch-based image retrieval,” *Machine Vision and Applications*,vol.29,no.7, pp.1083-1100,2018.

[6] S. Parui and A. Mittal, “Similarity-invariant sketch-based image retrieval in large databases,” *In European Conference on Computer Vision*, Springer, Cham, Zurich, Switzerland, pp. 398-414,2014.

[7] H.A. Abdulbaqi, G. Sulong and S.H. Hashem, “A sketch-based image retrieval: a review of literature,” *Journal of Theoretical and Applied Information Technology*, vol.63, no.1, pp.158-167,2014.

[8] F. Liu, X. Deng, C. Zou, Y.K. Lai, K. Chen *et al.*, "Scenesketcher-v2: fine-grained scene-level sketch-based image retrieval using adaptive gcns," *IEEE Transactions on Image Processing*, vol. 31, pp. 3737-3751, 2022,

[9] Y. Chen, Z. Zhang, Y. Wang, Y. Zhang, R. Feng *et al.*, “AE-Net: fine-grained sketch-based image retrieval via attention-enhanced network,” *Pattern Recognition*, vol. 122, pp.108291,2022.

[10] X. Zhang, M. Shen, X. Li and F. Feng, “A deformable cnn-based triplet model for fine-grained sketch-based image retrieval,” *Pattern Recognition*, vol.125, pp.108508,2022.

[11] H. Ren, Z. Zheng, Y. Wu, H. Lu, Y. Yang *et al.*, “ACNet: approaching-and-centralizing network for zero-shot sketch-based image retrieval,” arXiv preprint arXiv:2111.12757,2021.

[12] N. Kumar, R. Ahmed, V.B. Honnakasturi, S.S. Kamath and V. Mayya “Sketch-based image retrieval using convolutional neural networks based on feature adaptation and relevance feedback,” *In International Conference on Emerging Applications of Information Technology*, Springer, Singapore, pp. 103-113,2021.

[13] Q. Yu, J. Song, Y.Z. Song, T. Xiang and T.M. Hospedales, “Fine-grained instance-level sketch-

based image retrieval," *International Journal of Computer Vision*, vol.129, no.2, pp.484-500,2021.

[14] W. Zhou, J. Jia, C. Huang and Y. Cheng, "Web3d learning framework for 3d shape retrieval based on hybrid convolutional neural networks," *Tsinghua Science and Technology*, vol.25, no.1, pp.93-102,2019.

[15] U. Chaudhuri, B. Banerjee, A. Bhattacharya and M. Datcu, "CrossATNet-a novel cross-attention based framework for sketch-based image retrieval," *Image and Vision Computing*, vol.104, pp.104003,2020.

[16] R. Anisha, N. Anusha and G. Kavya, "Enhancing sketch-based image retrieval by cnn semantic re-ranking," *Annals of the Romanian Society for Cell Biology*, vol.25, no.4, pp.17812-17816,2021.

[17] L. Wang, X. Qian, X. Zhang and X. Hou, "Sketch-based image retrieval with multi-clustering re-ranking," *IEEE Transactions on Circuits and Systems for Video Technology*, vol.30, no.12, pp.4929-4943,2019.

[18] M. Ragab, A. Almuhammadi, R.F.Mansour and S. Kadry, "Natural language processing with deep learning enabled hybrid content retrieval model for digital library management," *Expert Systems*, pp.e13135,2022.

[19] A.K. Bhunia, Y. Yang, T.M. Hospedales, T. Xiang and Y.Z. Song, "Sketch less for more: on-the-fly fine-grained sketch-based image retrieval," *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA, pp. 9779-9788,2020.

[20] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, San Diego, CA, USA, vol.1, pp. 886-893, June 2005.

[21] R. Hu, M. Barnard and J. Collomosse, "Gradient field descriptor for sketch-based retrieval and localization," *2010 IEEE International Conference on Image Processing*, Hong Kong, China, pp. 1025-1028, Sep. 2010.

[22] J. M. Saavedra, "Sketch based image retrieval using a soft computation of the histogram of edge local orientations (S-HELO)," *2014 IEEE International Conference on Image Processing (ICIP)*, Paris, France, pp. 2998-3002, Oct. 2014.

[23] J. M. Saavedra and B. Bustos, "Sketch-based image retrieval using key shapes," *Multimedia Tools and Applications*, vol. 73, pp.2033–2062, 2014.

[24] L. Liu, F. Shen, Y. Shen, X. Liu and L. Shao, "Deep sketch hashing: fast free-hand sketch-based image retrieval," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, pp. 2298-2307, July 2017.

[25] F. Wang, L. Kang and Y. Li, "Sketch-based 3d shape retrieval using convolutional neural networks," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, pp. 1875-1883, June 2015.

[26] Q. Yu, Y. Yang, Y. Z. Song, T. Xiang and T. M. Hospedales, "Sketch-anet that beats humans," *International Journal of Computer Vision*, vol.122, pp.411-425,2017

[27] Y. Qi, Y. Z. Song, H. Zhang and J. Liu, "Sketch-based image retrieval via siamese convolutional neural network," *2016 IEEE International Conference on Image Processing*, Phoenix, AZ, USA, pp. 2460-2464, Sep. 2016.

[28] P. Sangkloy, N. Burnell, C. Ham and J. Hays, "The sketchy database: learning to retrieve badly drawn bunnies," *ACM Transactions on Graphics*, vol. 35, no. 4, pp. 1-12, Jul. 2016.

[29] Y. Song, J. Lei, B. Peng, K. Zheng, B. Yang *et al.*, "Edge-guided cross-domain learning with shape regression for sketch-based image retrieval," *IEEE Access*, vol.7, pp- 32393-32399, 2019.

[30] D.G. F. Pachecoa, J. Conesaa and N. Aleixos, "A new agent-based paradigm for recognition of free-hand sketches," *International Conference on Computational Science*, vol.1, no.1, pp.2013-2022, May 2010.

[31] S. Abbasi, F. Mokhtarian and J. Kittler, "Curvature scale space image in shape similarity retrieval," *Multimedia Systems*, vol.7, pp. 467–476, 1999.

[32] D. Zhang and G. Lu , "A comparative study of curvature scale space and fourier descriptors for shape-based image retrieval," *Journal of Visual Communication and Image Representation*,vol.14,no.1,pp.39-57,2003.

[33] V.A. Krylov and J.D.B. Nelson, "Stochastic extraction of elongated curvilinear structures with applications," *IEEE Transactions on Image Processing*, vol. 23, no. 12, pp. 5360-5373,2014.

[34] J. Xiao, Z.P. Tang, Y. Feng and Z. Xiao, "Sketch-based human motion retrieval via selected 2d geometric posture descriptor," *Signal Processing*, vol.113, pp.1-8, 2015.

[35] N. Prajapatil and G.S. Prajapti, "Sketch based image retrieval system for the web - a survey," *International Journal of Computer Science and Information Technologies*, vol. 6, no.4, pp. 3973-3979,2015.

[36] S. Tiwari, M.S. Shaikh and S. Gavhane, "Image retrieval by matching sketches and images," *International Journal of Engineering and Innovative Technology (IJEIT)*, vol. 3, no.10, pp.68-74,2014.

[37] S.R. Pawarl and K.R. Kandharkar, "Sketch based image retrieval system*," International Journal of Emerging Technology and Advanced Engineering*, vol. 4, no.10, pp.328-331,2014.

[38] PA. Gaidhani and S. B Bagal, "Survey paper on sketch based and content-based image retrieval," *International Journal of Science and Research (IJSR)*, vol.4, no.12, pp. 1-7, 2015.

[39] T. Bui and J. Collomosse, "Scalable sketch-based image retrieval using color gradient features," *2015 IEEE International Conference on Computer Vision Workshop (ICCVW)*, Santiago, Chile, pp. 1012-1019, Dec.2015.

[40] Y. H. S. Kumara and D. S. Gurub, "Retrieval of flower based on sketches," *Procedia Computer Science*, vol.46, pp. 1577-1584,2015.

[41] S. Shinde, N. Priya and N. Harpreetkaur, "Sketch based image retrieval system using wavelet transform," *International Journal of Innovative Research & Development*, vol. 2, no.4, pp.69-74,2013.

[42] K. M. Wilson and R. Shihab, "The enhanced sketch based image retrieval using semi-supervised biased maximum marginal analysis," International *Journal on Recent Trends in Engineering and Technology*, vol. 11, no. 1,pp.219, 2014.

[43] A.R. Moncy and S.P Sekhar, "Sketch-based image retrieval and enhancement by re-ranking and relevance feedback," *International Journal of Advanced Scientific Technologies, Engineering and Management Sciences*, vol.3, no.1, pp.12-15, 2017.

[44] K.Y. Tseng, Y. L. Lin, Y.H. Chen and W.H. Hsu, "Sketch-based image retrieval on mobile devices using compact hash bits," *MM '12: Proceedings of the 20th ACM international conference on Multimedia*, Nara Japan, pp. 913–916,2012.

[45] Y. Li and W. Li, "A survey of sketch-based image retrieval," *Machine Vision and Applications*, vol. 29, pp.1083–1100, 2018.

[46] Y. Song, J. Lei, B. Peng, K. Zheng, B. Yang *et al.*, "Edge-guided cross-domain learning with shape regression for sketch-based image retrieval," *IEEE Access*, vol.7, pp- 32393-32399, 2019.

[47] J. Sheng, F. Wang, B. Zhao, J. Jiang, Y. Yang *et al.*, "Sketch-based image retrieval using novel edge detector and feature descriptor," *Wireless Communications and Mobile Computing*, vol.2022, pp.1-12,2022.

[48] Q. Qi, Q. Huo, J. Wang, H. Sun, Y. Cao *et al.*, "Personalized sketch-based image retrieval by convolutional neural network and deep transfer learning," in *IEEE Access*, vol. 7, pp. 16537-16549, 2019.