International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V72I10P115
Volume 72 Issue 10, 149-158, October 2024

Original Article

QWO-IRV2 Deep Learning Model for Performing Retrieval of Multi Object Images

Sarva Naveen Kumar[1], Ch Sumanth Kumar[2]

[1,2]*Department of ECE, Gitam School of Technology, Gitam (Deemed to be University), Visakhapatnam, India.*
[1]*Department of ECE, MLR Institute of Technology, Hyderabad, India.*

[1]*Corresponding Author : navin9ap@gmail.com*

Received: 20 June 2024 Revised: 02 October 2024 Accepted: 14 October 2024 Published 25 October 2024

Abstract - *In the present scenario, data storage in the form of images is developing over the internet and effective techniques need to be implemented for retrieval of images. A large amount of data is available for numerous applications. Content-Based Image Retrieval (CBIR) is one of the important aspects that need to be considered. The goal of the paper is to retrieve multi object images from the database, which is a difficult issue for images that contain multiple objects. To overcome the problem, a deep learning technique along with an optimization method is developed for effective retrieval of multi object images. The Deep Convolutional Neural Network (DCNN) is designed. The network designed is ResNet 101 and Inception ResNetV2 (IRV2) for extraction of different types of features for the given input data. The features extracted depend on colour, shape, texture and so on. All the features are optimized using Quantum Whale Optimization (QWO). The final step is to estimate the similarity distance between the input query image and the data of images stored. Euclidean distance is utilized to identify the similarity index or level of similarity. The experimental evaluation is performed on the Corel database. The effectiveness and efficiency of the work are portrayed by utilizing metrics like recall, specificity, precision, and accuracy. The overall accuracy obtained by the proposed model is 98.33%. The suggested work QWO-IRV2 enhances image retrieval performance on the benchmark dataset in terms of accuracy, precision, and recall.*

Keywords - *Content based image retrieval, Convolutional neural networks, Inception RestNet v2, Quantum whale optimization, Corel dataset.*

1. Introduction

When no additional information is provided, CBIR systems are designed to enable users to obtain images associated with a semantic idea of their interest. A CBIR system often represents the photos in the repository as a multi-dimensional feature vector that is derived from several low-level descriptors, including colour, texture, and form. Next, a distance/similarity function established on the appropriate multi-dimensional feature space is used to quantify the perceptual similarity between two images. In recent years, many studies have focused on the implementation of CBIR systems, and several features have been proposed. [1] contains a comprehensive collection of commonly used picture descriptors. The "semantic gap," or the challenge of converting user intents into commonalities across low-level characteristics, is a significant issue with CBIR systems. A concept from classical information retrieval called relevance feedback has been widely employed to boost CBIR systems' effectiveness. This is made up of a system that analyses user input to improve the initial query iteratively. Every time the system performs an iteration, it retrieves and presents a selection of photographs and requests comments from the user about the relevance or non-relevance of each result to their search. This data is sent into an algorithm that modifies the search strategy and similarity metric to provide a fresh set of results. The resulting algorithm defines the CBIR method and determines the retrieval performance. The majority of CBIR algorithms in use today prioritize utilizing identified regions of interest over exploring the whole solution space. This implies that, based on certain semantic criteria, photos that are highly similar to samples that the user has already determined to be relevant are prioritized. This strategy makes sense when it comes to optimizing retrieval performance at a specific iteration but not when it comes to attaining peak performance during the whole relevant feedback process. Choosing the most instructive samples for evaluation is just as crucial as optimizing retrieval in a given iteration. It is possible for exploitation to produce a greater number of pertinent samples during a given iteration. Relevance data, however, will only be evaluated in the same locations in the subsequent iteration. Abuse of exploration, on the other hand, may result in poor retrieval efficiency that would take a long time to amortize.

Computer vision research has been greatly impacted by deep learning, and this effect has grown along with the advancement of deep learning, particularly in terms of Convolutional Neural Network (CNN) performance. Image classification issues on ImageNet datasets have led to improvements in CNN performance in recent years. Numerous notable CNN models, including Alexnet, VGG, Googlenet, and Resnet, have been proposed. As model depth increases, newly proposed models update the database of picture categorization jobs. In most of the CNN models, the pretraining is performed using ImageNet datasets and some of them are used for feature extraction in the images. Because of CNN's excellent feature encoding capabilities, network learning based on the ImageNet classification problem offers some invariance and may be implemented in a variety of application sectors with equivalent picture datasets.

In this paper the CBIR is processed using deep learning concept for the purpose of extracting the features of the images. Deep learning plays a crucial role in extracting the features. The proposed deep CNN model, i.e., the Inception ResNet V2 model helps in extracting the most essential features for the given input database. Further, the model involves meta heuristic algorithm for optimizing the obtained features. This paper was designed using two important concepts:

- Combination of two deep learning neural networks, Inception v2 and ResNet model, into Inception ResNet V2 model for extracting image features.
- A quantum whale optimization model is designed for selecting the optimized features. Finally, the Euclidean distance function is used to compute the distance between the optimized features and the feature database that is accessible in the dataset.

The paper is organized as follows. Section 2 discusses the existing models in CBIR. Section 3 discusses the proposed framework. The experimental results evaluated are showcased in Section 4. The discussion of total work has been concluded in Section 5.

2. Related work

In [2], the CBIR is performed by enhancing the process of feature extraction. Features are fed to a classifier, and the performance model is evaluated. The precision results are identified by using the classification model. The models used in the work are artificial neural networks, fuzzy neural networks, and navies Bayesian classifiers. The author in [3] concentrated on colour feature components for the retrieval of images. The fundamental concepts of colour are involved, and preference was shown for colour based features. The results obtained using the model are 84% rate of accuracy. In [4], a technique called multi trend structure descriptor is utilized by the author for feature extraction. The algorithm model helps in achieving lower-level features and information on local spatial structural features. The standard level datasets, such as corel-1k and corel-5k, are utilized for evaluating the results of the proposed model. The author concentrated on the time factor. The results show that 9.16 seconds were taken for retrieving the images. The accuracy achieved is around 80%. In [5] author presented an algorithm, i.e., Fast Image Retrieval (FIR), and implemented it on standard Corel datasets. The rate accuracy is 69%, which is lower and needs to improve.

The author in [6] proposed a Point of Interest Region (POIR) for feature extraction. This detection of the descriptor model, along with the SIFT model, shows an accuracy of 74.6%. The proposed approach used standard datasets, i.e., corel-1k and corel-5k. In [7] author presented a Convolution Neural Network (CNN) model for the extraction of features for performing the CBIR model. Along with CNN, the Euclidian Distance (ED) model is designed to measure the similarity of images and the exactness of images.

This model is evaluated in three different datasets and achieved the results. The results obtained using the Corel dataset is 95%, using the Caltech dataset 96.5% and for the Li dataset 88%. The author in [8] proposed a SIFT based model for the extraction of features and bacteria foraging optimization for achieving an optimized set of features. The optimization techniques provide a selected list of features that are useful for retrieving the required image. The cost and time factors will be reduced. The performance of the model is evaluated on images like animals, faces and flowers. In [9] author introduced a recovery of images model based on Log-Gabor channels. The model is been tested on datasets like Vistex and OT-scene.

The author in [10] proposed an Edge Histogram Descriptor (EDH) model for removing the substances in the images for retrieval of images. This framework gives decent output results. The idea of obtaining the results helps in incorporating the MPEG-7 images. In [11] the author considered some features like mean, standard deviation and homogeneity for CBIR. The facial images are considered and evaluated. The neighbourhood image identification model is used to investigate the face of local points.Some of the other different models proposed by various authors are Fuzzy logic [12], Clustering based techniques [13], Self-organizing map model [14], and Genetic optimization model [15].

The search models for image retrieval concentrate on the surrounding samples in the search space. Various feature extraction-based models designed by different others using the Corel dataset are shown in Table1. From Table 1, it is observed that retrieval of images from the dataset achieved less accuracy, and hence, improving accuracy is an important task. Multi objects retrieval needs to be concentrated as the image consists of different objects, and many researchers concentrated on single object retrieval.

Table 1. Existing methods of CBIR

Author	Dataset utilized	Approach	Result
X.Y Wang et al. [16]	Corel	Fusion of color and texture features for CBIR	Accuracy= 61.3%
J.M. Guo et al. [17]	Corel	Block truncation coding (BTC) features for CBIR	Accuracy= 79.7%
H. Shao et al. [18]	M	Extraction of color features for image retrieval	Accuracy= 89.6%
Y. Liu et al. [19]	Corel	Decision Tree Learning for region-based image retrieval	Accuracy= 76.8%
M.M Islam et al. [20]	Corel	Colour-based vector quantization for CBIR	Accuracy= 97.6%
Z. Jiexian et al. [21]	MPEG-7	Coherence vector algorithm for CBIR	Accuracy= 97%
A. Nazir et al. [22]	Corel 1K	HSV colour histogram, DWT for CBIR	Precision= 73.5%
R. Ashraf et al. [23]	Corel 1000	Multiple feature extraction using multimedia data for CBIR	Precision= 87.5%
Y. Mistry et al. [24]	Corel	Hybrid feature and distance parameters for CBIR	Precision= 87.5%

3. Proposed Methodology

A method called Content-Based Picture Retrieval (CBIR) uses visual contents, or picture characteristics, to search databases. To locate similar pictures, CBIR compares the query image to an image stored in the database or uses the input text as an additional input. To get the feature space that is recorded in the database of image features, the CBIR works by first extracting features from the picture in the database. The features will be compared one by one with the features in the database after the query picture is provided. The picture is identical to the one that must be obtained at the shortest distance. In the case of input text, all the images corresponding to the text are identified. Every image available in the database has it text label for retrieving. The block diagram of the proposed framework is shown in Figure 1.

3.1. Input Data

To evaluate the performance of the Deep Neural Networks (DNN) for CBIR a standard dataset is utilized. In this paper, the Corel database is utilized [25]. The collection of images in the database is 1000. The data is divided into twenty sub-categories with images like birds, bicycles, buses, cats, cars, horses, sheep, trains etc. Each category has 50 number of images. The images available in the dataset are preprocessed before performing the feature extraction and optimization process.

3.2. Stage of Preprocessing

Preprocessing is done to extract the correct region of the picture that corresponds to the analysis of image retrieval using various techniques, such as border detection, and to eliminate distortions and other undesired aspects when processing the image. Preprocessing includes normalization, border detection, scaling the picture, and eliminating undesired noise.

Preprocessing involves several steps for the picture, including resizing, detection of borders, and normalization. In this paper, primarily, all the images in the database are resized to 227×227 and any unwanted noise caused by the system while loading the data has been removed. Further, the feature extraction process is executed.

3.3. Feature Extraction

The important step in CBIR is feature extraction. The process of feature extraction is performed both on the input database and input query image. To extract the features, a deep CNN model, i.e., ResNet 101 and Inception ResNet V2 (IRV2) model are designed. Using the deep CNN model, the shape features, colour features, spatial features and textural features are extracted. These features help in the representation of images and retrieval of images. The author in [26] combined Inception and ResNet architectures to form a hybrid model. The designed network consists of 166 deep layers, which leads to achieving better features and a better rate of accuracy in retrieving images. A CNN architecture called Inception ResNet v2 expands upon the Inception family of architectures by adding residual connections, which take the place of the Inception design's filter concatenation stage. As opposed to learning unreferenced functions residual connections, a particular kind of skip connection learns residual functions with the help of the input layers. The architecture of the proposed model with different layers is shown in Figure 2.

3.4. Optimization of Features

The features extracted are fed to the optimizer. The optimization technique helps in selecting the optimized features which makes the retrieving process faster and accurate. A Quantum Whale Optimization (QWO) is utilized in this work. The WOA takes the lead in the search by selecting a random whale location or by using the best solution found. This helps WOA behave in an exploratory manner and keeps the system from being trapped in local minima. The binary form of the feature vectors is used in feature selection based on WOA. Because of this, a large population is needed for an efficient search, which takes longer to find the best answer through exploration and exploitation. Our proposal is the Quantum-based WOA (QWOA), which is based on quantum computing principles and takes advantage of a probabilistic representation of the Q-bits to improve population diversity while managing the time and space complexity of the classical WOA and choosing an ideal number of features. The selection model of best features is done by utilizing the bubble-net process of the whale, as shown in Figure 3.

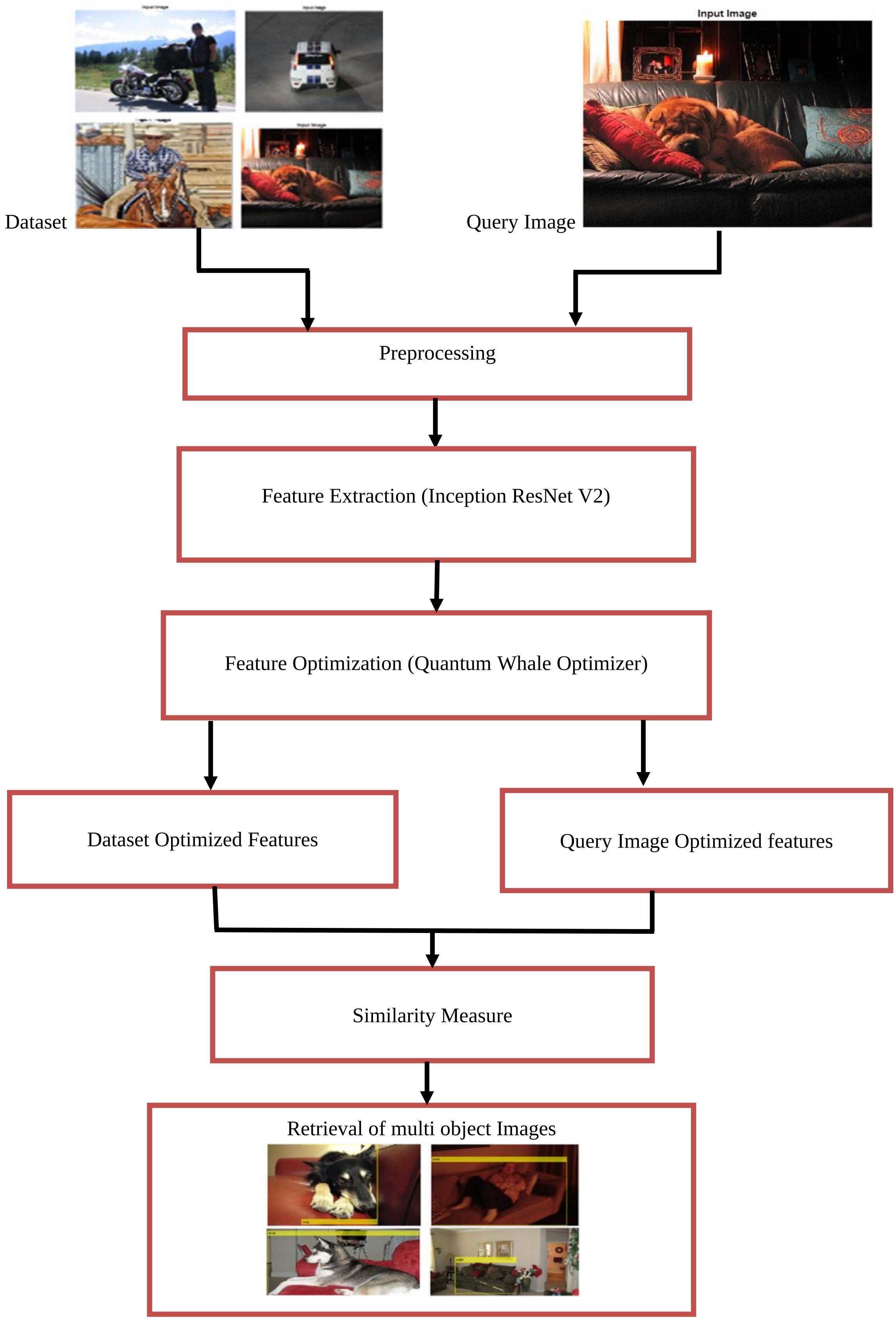


Fig. 1 Block diagram of the proposed methodology

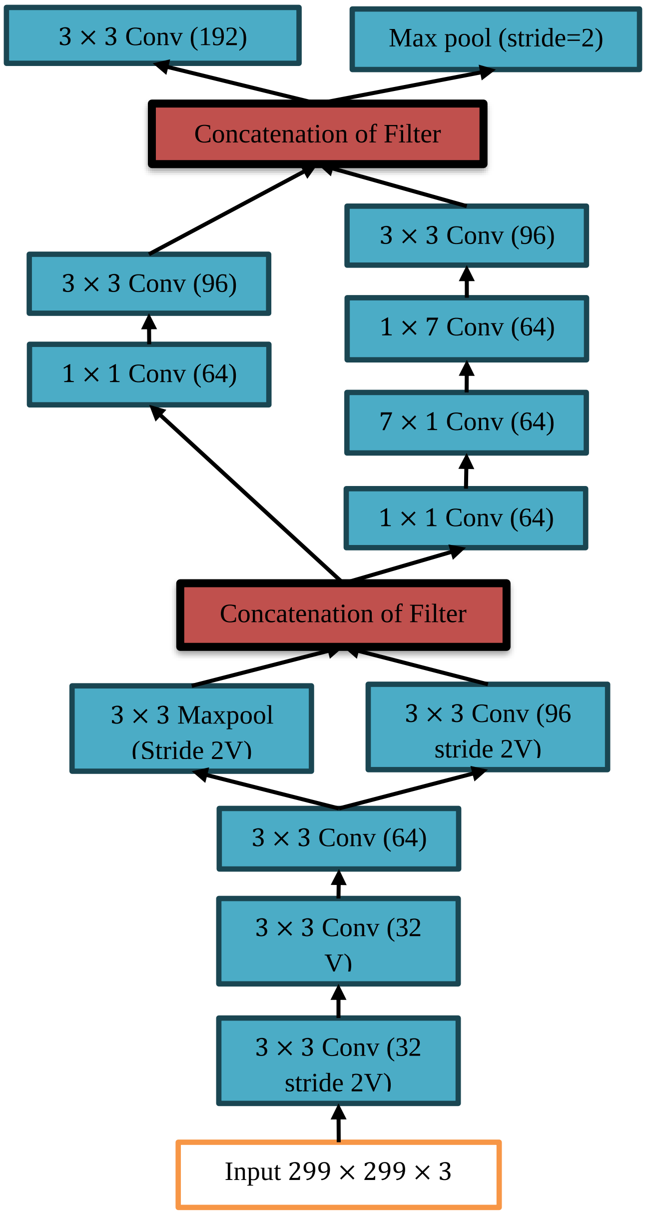


Fig. 2 Inception ResNet V2 design flow

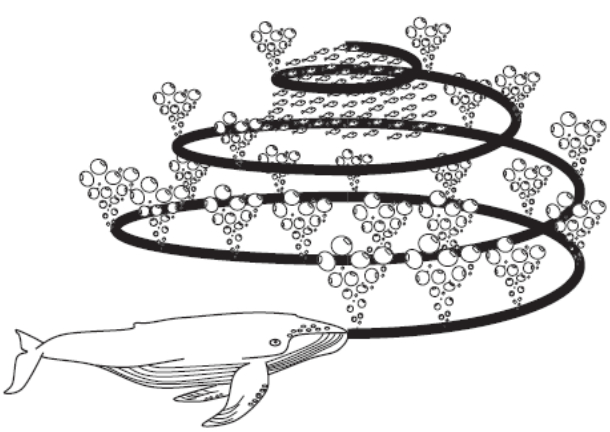


Fig. 3 Bubble-net process of whale

The process of selecting the best features using QWO is explained in steps.

Step 1: In a population of size n, each whale position is expressed in terms of d-dimensional Q-bits, which correspond to a single individual.

Step 2: The tournament picking technique is used to pick a randomly chosen person to lead the search during exploration.

Step 3: The mutant random individual and the mutated present individual undergo a crossover.

Step 4: Using the crossover between the best person and its modified counterpart, the shrinking approach is used. The spiral movement is implemented with the Equation (1) below.

$$D = W^{*} - W^{i} \tag{1}$$

($W^{*}$is mutated and crossed over with the current whale, W^{i} to give the updated position of the current whale)

Step 5: The modified spiralling position is calculated as follows:

$$W^{i} = D.e^{pl}.\cos(2\pi l) + W^{*} \tag{2}$$

(Here, p indicates how the logarithmic spiral is shaped. The interval [-1,1] is used to select the value of l. The dot operator represents the element-by-element multiplication.

Step 6: Each binary feature vector's quality of fit was assessed using the fitness function F, which was calculated as

$$F = (err) + (k) \tag{3}$$

Where α and β give the weight to err and k, respectively.

Step 7: The feature subset that minimizes both the feature subset size (k) and the classification error (err) is the one that is chosen.

Step 8: Up until the termination condition is fulfilled, the entire population is updated using the shrinking, spiralling, and exploring procedures based on the fitness values determined at each generation.

The features obtained using the IRV2 model are optimized by following the above steps. The optimized features of both inputs are fed for similarity check.

3.5. Similarity Measure

The similarity measure is performed using Euclidian Distance (ED) model. The similarity determines the difference between the features obtained for the dataset and the features obtained for the query image. The distance between the optimized features is calculated using the ED model. If the distance evaluated between the two images is small, then the similarity is greater for the query image and the generated image. Equation (4) shown below is used to compare the distance [27].

$$((x,y),(a,b))=\sqrt{(x-a)2+(y-b)2} \quad (4)$$

The terms x *and* y in Equation 1 refer to the two dimensions of a picture, and the images that are available according to the dimension in the dataset are termed as a *and* b.

Process of retrieving the image:

Step 1: All the feature vector in the image is labelled for all the classes.

Step 2: The nominal values of the classes are converted to numerical values.

Step 3: The dataset is divided into training and testing with a ratio of 80:20.

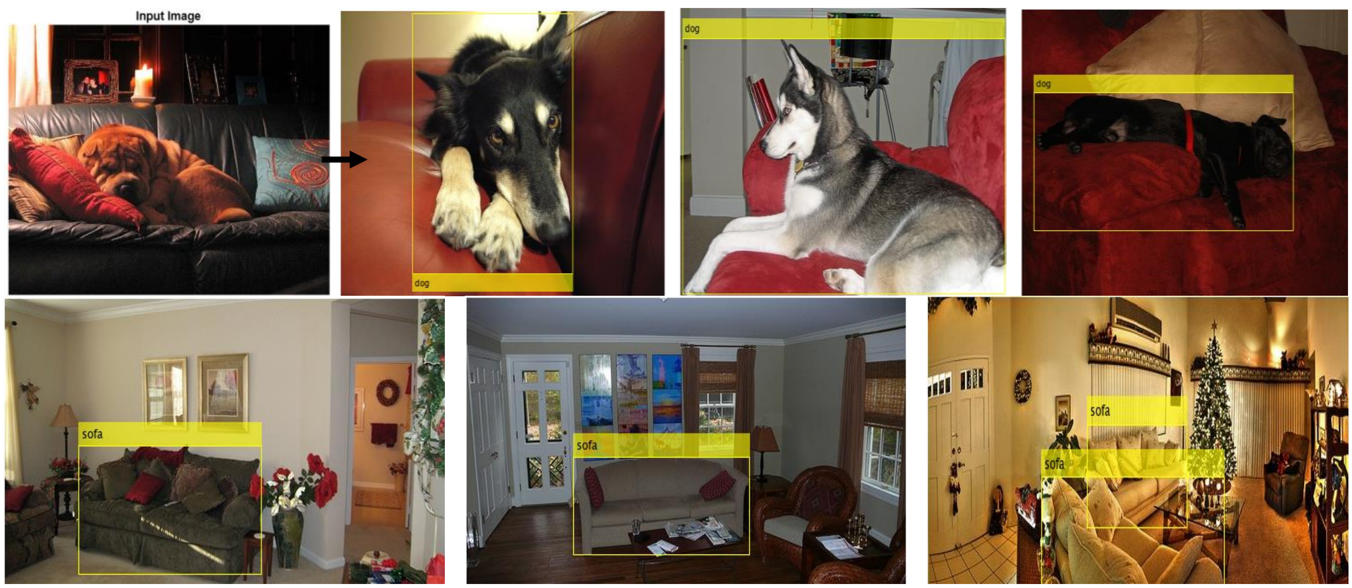
Step 4: Calculate the distance between the feature vectors of all training and testing data and achieve a smaller distance.

Step 5: Evaluate the mean average precision of all query images.

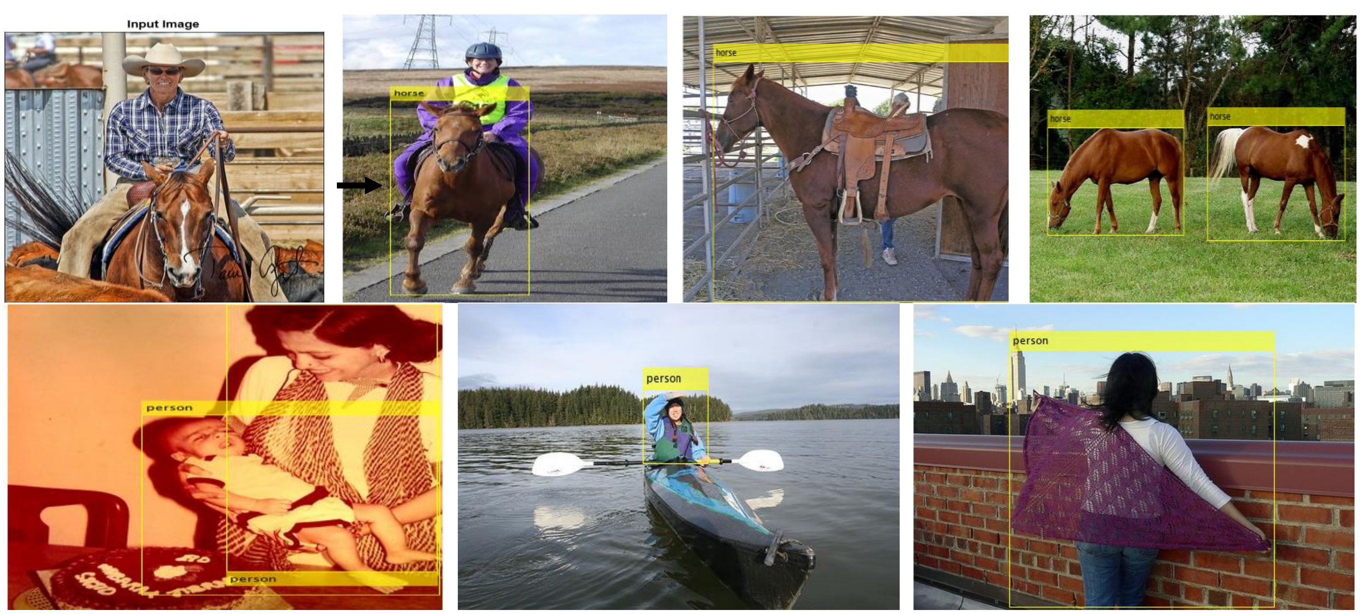
Step 6: Finally, display the images related to the query image.

4. Results and Discussion

The validation of the proposed IRV2- Quantum whale optimization is discussed in this section. The overall assessment is to show the performance of the proposed model utilizing the Corel database. This section includes a comparison of results achieved with existing methods.



(a)



(b)

Fig. 4 (a) Input image: Retrived images with dog and sofa (b) Input image: Retrieved horse and human images

The involvement of deep neural networks and optimization techniques in CBIR achieves good results. The set of images in the Corel database consists of different wide classes of images. In this wide range of images, an input query is given to retrieve the images as per the requirement.

The input query can be an image or text. In this paper, the discussion is about retrieving multi object images. The images retrieved using the proposed model are shown in Figure 4. In Figure 4(a), the given input images consist of a dog and sofa as major contents in the images.

The outputs retrieved consist of both dog images and sofa-based images. In Figure 4(b) input image consists of a person and a horse. Hence, the retrieval of images with horse and person is obtained as output. The quantum whale optimizes the features that are obtained using the inception restnet v2 model.

The existing techniques, like quantum cuckoo search based feature selection and other methods, are compared with the proposed model. The performance of the model designed is evaluated using the following metrics.

$$Sensitivity\ (Se) = \frac{TP}{TP+FN} \tag{5}$$

$$Specificity\ (Sp) = \frac{TN}{TN+FP} \tag{6}$$

$$Accuracy\ (A) = \frac{TP+TN}{TP+TN+FP+FN} \tag{7}$$

$$Precision\ (P) = \frac{Number\ of\ retrived\ relevant\ images}{No\ of\ relevent\ images} \tag{8}$$

*TP- True Positive; TN- True Negative; FN- False Negative; FP- False Positive.

The values evaluated using only deep learning techniques, i.e., ResNet-101 and a Euclidian distance model, are shown in Table 1. The optimization technique is not involved, and the overall retrieval accuracy obtained is around 94.71%. After combing deep networks with optimization techniques results are evaluated further.

The results evaluated are shown in Table 2. The values are obtained by applying an optimization approach. In this case, Euclidian distance and QWO-ResNet 101 are used to analyze the parameters. The overall rate of accuracy obtained in the retrieval of images is 96.98%. In Table 3, the results obtained using QWO-IRV2 are shown.

The overall accuracy achieved using the technique is 98.33%. Table 4 displays the values of the comparison between the current methodologies and the overall accuracy and other parameters.

In Figure 5, the graphical analysis is displayed. Because of the precise clustering model, which is depicted in Figure 5, the QWO-IRV2 model outperformed the existing QCSA-Alexanet [27] and other models in terms of accuracy rate and other characteristics as well.

Table 1. Results obtained using ResNet 101

Class Name	Se/Recall (%)	Precision (%)	Specificity (%)	Error Percent	Accuracy (%)
‘Aeroplane’	94.13	67.33	94.18	5.50	94.5
‘Bicycle’	91.49	66.42	92.3	5.96	94.03
‘Bird’	93.44	73.30	93.84	5.45	94.55
‘Boat’	95.55	64.67	95.88	5.15	94.84
‘Bottle’	93.68	69.88	93.98	5.50	94.49
‘Bus’	95.93	67.04	96.12	4.98	95.01
‘Car’	95.73	83.45	95.98	5.12	94.88
‘Cat’	94.00	72.38	94.42	5.31	94.68
‘Chair’	94.71	75.87	94.82	5.61	94.38
‘Cow’	94.69	63.54	95.21	4.82	95.17
‘Table’	95.43	65.06	95.75	5.96	94.03
‘Dog’	95.72	74.07	96.10	5.88	94.11
‘Horse’	96.96	71.02	97.12	5.17	94.82
‘Motor-bike’	93.56	68.18	93.86	5.45	94.55
‘Person’	94.56	92.59	94.78	5.34	94.65
‘Potted-plant’	96.19	71.95	96.48	4.55	95.44
‘Sheep’	94.52	60.37	94.86	4.71	95.28
‘Sofa’	93.30	67.39	93.60	5.36	94.63
‘Train’	95.771	71.48	95.91	4.50	95.49
‘TV monitor’	95.322	69.32	95.62	5.39	94.60
‘Overall Percentage	94.738	70.76	94.73	5.29	94.71

Table 2. Results obtained using QWO- ResNet-101

Class Name	Se/Recall	Precision	Specificity	Error Percent	Accuracy
‘Aeroplane’	97.54	71.75	98.12	3.93	96.06
‘Bicycle’	97.72	76.48	98.34	2.79	97.20
‘Bird’	98.32	82.29	98.67	2.68	97.31
‘Boat’	96.41	69.16	96.82	3.47	96.52
‘Bottle’	97.32	74.79	97.65	3.82	96.17
‘Bus’	95.97	73.39	96.20	2.95	97.04
‘Car’	96.51	89.75	96.89	2.87	97.12
‘Cat’	96.27	79.11	96.52	3.19	96.8
‘Chair’	96.75	85.58	97.15	2.57	97.42
‘Cow’	97.072	69.22	97.40	2.95	97.04
‘Table’	95.71	71.32	96.08	3.47	96.53
‘Dog’	96.858	81.38	97.18	3.31	96.69
‘Horse’	96.842	76.20	97.12	3.41	96.58
‘Motor-bike’	97.307	76.54	97.90	2.87	97.12
‘Person’	97.177	95.49	97.80	3.03	96.96
‘Potted plant’	97.912	78.95	98.30	2.63	97.37
‘Sheep’	97.461	65.05	97.92	2.98	97.01
‘Sofa’	97.237	75.79	97.83	2.82	97.18
‘Train’	97.457	74.97	97.87	3.44	96.55
‘TV monitor’	97.295	73.92	97.54	3.79	96.23
‘Overall Percentage	97.10	77.60	97.28	3.15	96.98

Table 3. Optimized QWO- IRV2

Class Name	Se/Recall	Precision	Specificity	Error Percent	Accuracy
‘Aeroplane’	98.491	79.66	98.86	2.09	97.91
‘Bicycle’	96.809	79.92	97.40	2.14	97.85
‘Bird’	98.832	90.24	98.96	1.22	98.78
‘Boat’	98.891	75.15	98.92	2.16	97.83
‘Bottle’	99.042	83.57	99.26	1.84	98.15
‘Bus’	97.354	77.12	97.67	2.22	97.77
‘Car’	97.937	91.44	98.12	2.22	97.77
‘Cat’	99.069	89.31	99.26	1.22	98.78
‘Chair’	98.544	92.41	98.76	1.17	98.83
‘Cow’	99.171	76.7	99.37	1.63	98.37
‘Table’	98.802	83.09	98.96	1.35	98.64
‘Dog’	97.996	86.79	98.18	2.03	97.96
‘Horse’	98.759	89.34	98.92	1.06	98.94
‘Motor-bike’	99.524	88.96	99.58	0.92	99.08
‘Person’	98.412	97.65	98.65	1.59	98.4
‘Potted plant’	98.235	82.26	98.65	2.01	97.99
‘Sheep’	99.519	78.9	99.58	0.94	99.05
‘Sofa’	97.967	78.65	98.16	2.25	97.75
‘Train’	97.175	83.26	97.36	1.78	98.2
‘TV monitor’	98.865	85.26	98.98	1.46	98.53
‘Overall Percentage	98.47	88.48	98.87	1.66	98.33

Table 4. Comparison of the proposed method with existing methods

	CART- Decision Tree [28]	CNN-ED [29]	QCSA-Alexanet [27]	QWO-ResNet 101	Proposed QWO-IRV2
Recall (%)	91.24	94.99	97.13	97.10	98.47
Precision (%)	75.0	71.46	77.37	77.60	88.48
Specificity (%)	97.25	94.85	97.13	97.28	98.87
Accuracy (%)	82.3	94.98	96.91	96.98	98.33

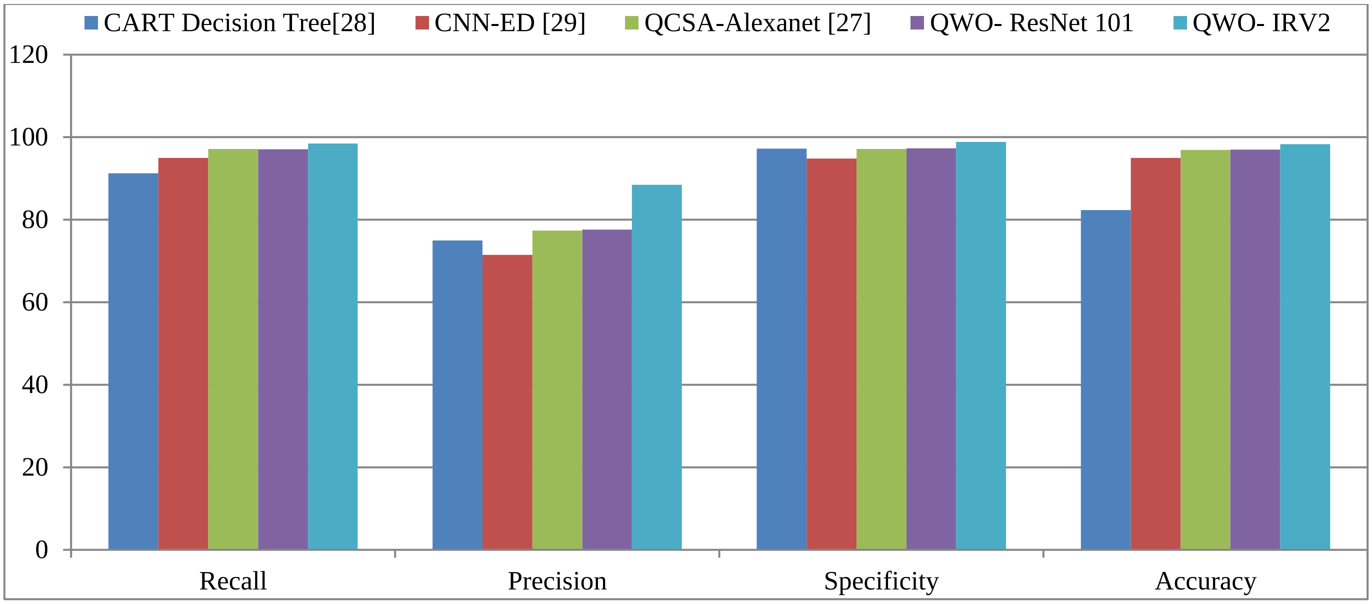


Fig. 5 Comparison of different techniques

5. Conclusion

The work aims to retrieve multi object-based images for the given query image. The proposed QWO-IRV2 model successfully retrieved a wide range of images based on the content. The designed work combines a deep learning model and optimization technique to improve the rate of accuracy in retrieving the images. The crucial problem for CBIR is the extraction of features. In this paper, the features are extracted using a hybrid deep learning model, i.e., Inception ResNet V2. Multiple features are been extracted and these features are optimized using QWO. The selection of the best features helps to improve the progress of work in terms of space and time. Finally, to retrieve the images Euclidean distance similarity measure is processed. Further enhancement can be performed by changing the optimization model, which helps in fine-tuning features and improves the rate of accuracy. The overall accuracy obtained using the proposed methodology is 98.333%.

References

[1] David Dagan Feng, Wan-Chi Siu, and Hong-Jiang Zhang, *Multimedia Information Retrieval and Management, Technological Fundamentals and Applications*, 1st ed., Springer Berlin, Heidelberg, pp. 1-476, 2003. [CrossRef] [Google Scholar] [Publisher Link]

[2] Manpreet Kaur, and Sakshi Dhingra, "Comparative Analysis of Image Classification Techniques Using Statistical Features in CBIR Systems," *2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, Palladam, India, pp. 265-270, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[3] Naushad Varish, and Arup Kumar Pal, "Content Based Image Retrieval Using Statistical Features of Color Histogram," *2015 3rd International Conference on Signal Processing, Communication and Networking*, Chennai, India, pp. 1-6, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[4] Meng Zhao, Huaxiang Zhang, and Jiande Sun, "A Novel Image Retrieval Method Based on Multi-Trend Structure Descriptor," *Journal of Visual Communication and Image Representation*, vol. 38, pp. 73-81, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[5] Liang Zheng et al., "Fast Image Retrieval: Query Pruning and Early Termination," *IEEE Transactions on Multimedia*, vol. 17, no. 5, pp. 648-659, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[6] K. Mikolajczyk, and C. Schmid, "A Performance Evaluation of Local Descriptors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1615-1630, 2005. [CrossRef] [Google Scholar] [Publisher Link]

[7] Amjad Shah et al., "Improving CBIR Accuracy Using Convolutional Neural Network for Feature Extraction," *2017 13th International Conference on Emerging Technologies*, Islamabad, Pakistan, pp. 1-5, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[8] Aasia Ali, and Sanjay Sharma, "Content Based Image Retrieval Using Feature Extraction with Machine Learning," *2017 International Conference on Intelligent Computing and Control Systems*, Madurai, India, pp. 1048-1053, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[9] Ekta Walia, and Vishal Verma, "Boosting Local Texture Descriptors with Log-Gabor Filters Response for Improved Image Retrieval," *International Journal of Multimedia Information Retrieval*, vol. 5, pp. 173-184, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[10] Pradnya Vikhar, and Pravin Karde, "Improved CBIR System Using Edge Histogram Descriptor (EHD) and Support Vector Machine (SVM)," *2016 International Conference on ICT in Business Industry & Government*, Indore, India, pp. 1-5, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[11] C. Benavides et al., "Face Classification by Local Texture Analisys through CBIR and SURF Points," *IEEE Latin America Transactions*, vol. 14, no. 5, pp. 2418-2434, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[12] Chih-Yi Chiu, Hsin-Chih Lin, and Shi-Nine Yang, "A Fuzzy Logic CBIR System," *The 12th IEEE International Conference on Fuzzy Systems, 2003. FUZZ '03*, St. Louis, MO, USA, pp. 1171-1176, 2003. [CrossRef] [Google Scholar] [Publisher Link]

[13] Yixin Chen, J.Z. Wang, and R. Krovetz, "CLUE: Cluster-Based Retrieval of Images by Unsupervised Learning," *IEEE Transactions on Image Processing*, vol. 14, no. 8, pp. 1187-1201, 2005. [CrossRef] [Google Scholar] [Publisher Link]

[14] J. Laaksonen, M. Koskela, and E. Oja, "PicSOM-Self-Organizing Image Retrieval with MPEG-7 Content Descriptors," *IEEE Transactions on Neural Networks*, vol. 13, no. 4, pp. 841-853, 2002. [CrossRef] [Google Scholar] [Publisher Link]

[15] O. Cordón et al., "A Review on the Application of Evolutionary Computation to Information Retrieval," *International Journal of Approximate Reasoning*, vol. 34, no. 2-3, pp. 241-264, 2003. [CrossRef] [Google Scholar] [Publisher Link]

[16] Xiang-Yang Wang, Bei-Bei Zhang, and Hong-Ying Yang, "Content-Based Image Retrieval by Integrating Color and Texture Features," *Multimedia Tools and Applications*, vol. 68, pp. 545-569, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[17] Jing-Ming Guo, Heri Prasetyo, and Jen-Ho Chen, "Content-Based Image Retrieval Using Error Diffusion Block Truncation Coding Features," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 3, pp. 466-481, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[18] Hong Shao et al., "Image Retrieval Based on MPEG-7 Dominant Color Descriptor," *2008 The 9th International Conference for Young Computer Scientists*, Hunan, China, pp. 753-757, 2008. [CrossRef] [Google Scholar] [Publisher Link]

[19] Ying Liu, Dengsheng Zhang, and Guojun Lu, "Region-Based Image Retrieval with High-Level Semantics Using Decision Tree Learning," *Pattern Recognition*, vol. 41, no. 8, pp. 2554-2570, 2008. [CrossRef] [Google Scholar] [Publisher Link]

[20] Md Monirul Islam, Dengsheng Zhang, and Guojun Lu, "Automatic Categorization of Image Regions Using Dominant Color Based Vector Quantization," *2008 Digital Image Computing: Techniques and Applications*, Canberra, ACT, Australia, pp. 191-198, 2008. [CrossRef] [Google Scholar] [Publisher Link]

[21] Zeng Jiexian, Liu Xiupeng, and Fei Yu, "Multiscale Distance Coherence Vector Algorithm for Content-Based Image Retrieval," *The Scientific World Journal*, vol. 2014, no. 1, pp. 1-13, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[22] Atif Nazir et al., "Content Based Image Retrieval System by Using HSV Color Histogram, Discrete Wavelet Transform and Edge Histogram Descriptor," *2018 International Conference on Computing, Mathematics and Engineering Technologies*, Sukkur, Pakistan, pp. 1-6, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[23] Rehan Ashraf et al., "MDCBIR-MF: Multimedia Data for Content-Based Image Retrieval by Using Multiple Features," *Multimedia Tools and Applications*, vol. 79, pp. 8553-8579, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[24] Yogita Mistry, D.T. Ingole, and M.D. Ingole, "Content Based Image Retrieval Using Hybrid Features and Various Distance Metric," *Journal of Electrical Systems and Information Technology*, vol. 5, no. 3, pp. 878-888, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[25] Jia Li, and J.Z. Wang, "Automatic Linguistic Indexing of Pictures by a Statistical Modeling Approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1075-1088, 2003. [CrossRef] [Google Scholar] [Publisher Link]

[26] Christian Szegedy et al., "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31, no. 1, pp. 4278-4284, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[27] Sarva Naveen Kumar, and Ch. Sumanth Kumar, "Fusion of CNN-QCSO for Content Based Image Retrieval," *Journal of Advances in Information Technology*, vol. 14, no. 4, pp. 668-673, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[28] Gnanasigamony Wiselin Jiji, and Peter Savariraj Johnson Durai Raj, "Content-Based Image Retrieval in Dermatology Using Intelligent Technique," *IET Image Process*, vol. 9, no. 4, pp. 306-317, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[29] Maria Tzelepi, and Anastasios Tefas, "Fully Unsupervised Optimization of CNN Features Towards Content Based Image Retrieval," *2018 IEEE 13th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP)*, Aristi Village, Greece, pp. 1-5, 2018. [CrossRef] [Google Scholar] [Publisher Link]