VOL. 18, NO. 3, FEBRUARY 2023 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

QUERY EXTRACTION BASED ON ENCRYPTED FEATURES FOR CBIR FROM CLOUD

D. Ravibabu[1], Kesaboina Sushma[2], R. Shiva Shankar[1] and K. V. S. S. R. Murthy[2]
[1]Department of CSE, Sagi Rama Krishnam Raju Engineering College, Bhimavaram, Andhrapradesh, India
[2](CST), Department of CSE, Sagi Rama Krishnam Raju Engineering College, Bhimavaram, Andhrapradesh, India
E-Mail: ravibabu.devareddi@gmail.com

ABSTRACT

Due to the increasing use of mobile devices, content-based image retrieval is gaining massive popularity as mobile devices run on batteries. It cannot perform heavy image processing computation, so mobile users can extract just image features and offload them to the cloud. The cloud will perform content-based image searches and return a search result. COREL10K images dataset was selected, and VGG16, RESNET, DENSENET, and MOBILE NET algorithms were applied. Finally, the method with the best performance is preferred. Also, PAHE (Packed Additive Homomorphic Encryption) algorithm is selected to encrypt the feature vectors, and image similarity calculations are done with Secure Multi-party Computation (SMC). The following parameters are used for each algorithm to calculate accuracy, precision, recall, Jaccard index, error rates, and F1-score. The algorithms are then compared and selected the best algorithm. The best model obtained is VGG16, with the highest accuracy. The procedure can help Mobile users perform content-based image searches by extracting features and then uploading them to the cloud.

Keywords: packed additive homomorphic encryption (pahe), vgg16, resnet, densenet, mobile net, secure multi-party computation (smc), machine learning (ML).

1. INTRODUCTION

Images and graphics are among the most important forms of media for human communication because they provide a wealth of information that helps individuals better comprehend their surroundings. More photographs are accessible to the public because of the fast development of digital imaging methods and the internet [1]. To search through and retrieve certain photographs from among many digital photos, a computer system called an "image retrieval system" is required. There are two distinct categories of picture retrieval systems: those that rely on textual information and those that use contextual information. An image retrieval approach known as CBIR employs visual cues such as the colour, shape, and texture of a picture to find the most relevant image from a vast collection of images. These visual characteristics are retrieved mechanically or without human intervention. A recent trend in computer vision [2] has been the elevation of image search to a central role.

Facial recognition, image recognition, and video analysis are all AI applications that have become the essential formats that use the standard type of neural network architecture. In every image retrieval approach, the similarity calculation between pictures is critical. Ideally, the approach for computing similarity scores between two pictures should be discriminative, robust, and efficient [3].

When searching for images, content-based image retrieval may employ feature vectors to describe the picture's information. Different methods may be used to find similarity matching from the photos. Feature matching plays a significant role while finding the similarity matching between the images.

It is necessary to perform several manipulations on the sampled picture before any characteristics may be extracted. Features relevant to categorising and recognising pictures are then extracted using feature extraction methods [4]. For example, the picture is stored as matrices after the raw data is converted into numerical form through feature extraction. Using that matrix value, they compare the images and apply the max pooling. Here, both images have similar measures like colour and shape, etc. Histograms, sift, and so on these are some of the feature-extracting algorithms.

Textual annotation of pictures is becoming more problematic and wasteful for the representation and retrieval of images as digital photography equipment grows increasingly widespread. In the 1980s, CBIR began to emerge [1]. CBIR systems, such as QIBC, MARS, and Virage have been effectively produced by researchers over the last three decades [5]. Photobooks and FIDS systems have also been successfully constructed. Depending on all those, how good can image retrieval be done using different approaches were going to learn.

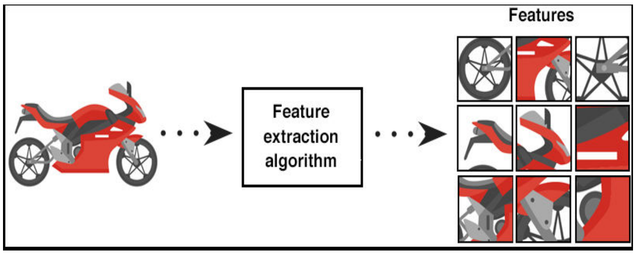

Figure-1. Feature extraction from the image.

Figure-1 illustrates the process of identifying and extracting characteristics from photos. Images were exhibited in Figure-2, and the design of the Text Image Retrieval System was illustrated in Figure-3.

ARPN Journal of Engineering and Applied Sciences

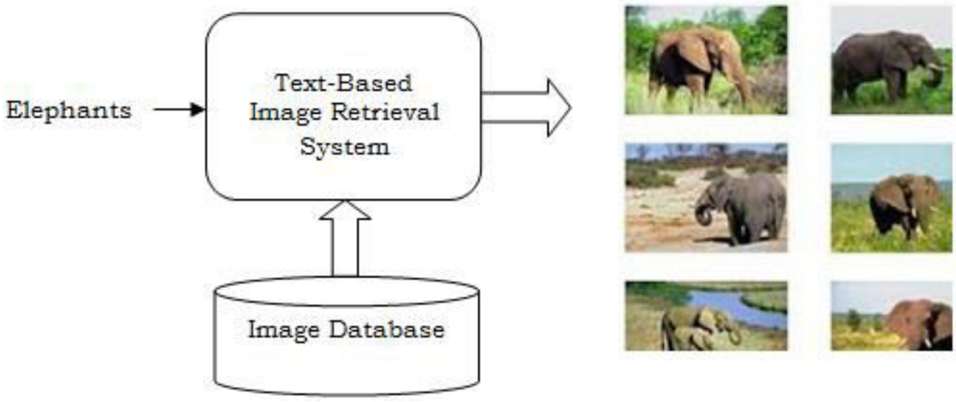

Figure-2. Basic Image Retrieval System.

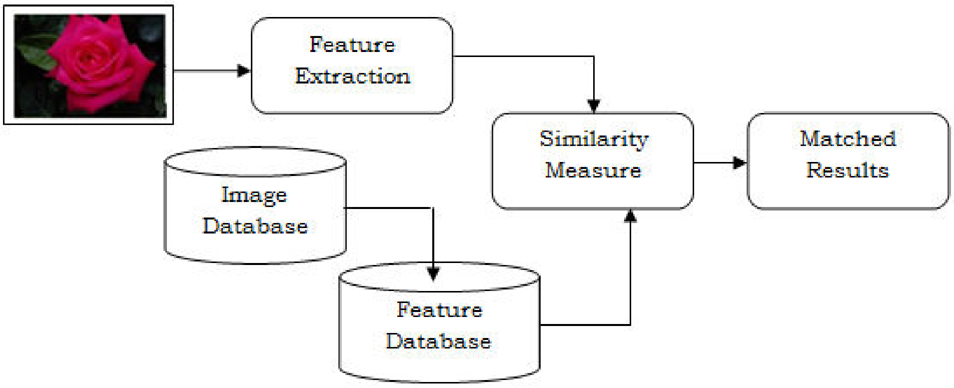

Figure-3. Text-based retrieval system.

2. RELATED WORK

An image feature detector named Micro Structure Descriptor was suggested by G H Liu *et al.* [1] (MSD). It's based on the underlying colours and edge orientations of the microstructure. Using Corel databases containing 15,000 natural photos, the suggested technique is evaluated for its effectiveness. Compared to a representative feature, it performs better when it comes to indexing full-color photos, but it has lesser dimensionality overall. Images may be retrieved utilising Local Tetra Patterns (LTrPs) for content-based image retrieval, as suggested by S Murala and R Maheshwari [5] (CBIR). The suggested approach encodes the connection between the referred pixel and its surrounding pixels. For horizontal and vertical directions of diagonal pixels, they compare the retrieval result with the usual Local Binary Patterns(LBP) and find an improvement of 70.34 percent to 75,9 percent using the Corel 10,000 databases.

A multi-trend structure descriptor (MTSD) image feature representation is used by M Zhao *et al.* [6]. It is reasonable to think of local structures as the foundational components of picture analysis. Results show that MTSD outperforms Textron co-occurrence matrix, multi-Textron histogram, microstructure descriptor, and saliency structure histogram.

Additional characteristics were included in the form of quantization (colour, edge orientation, and intensity map). An image feature descriptor called Local Bit Plane Decoded Pattern (LBDP) has been developed by S R Dubey [7] for the indexing and retrieval of biological pictures. The effectiveness of LBDP is evaluated using average accuracy and recall in biomedical image retrieval. Emphysema-CT, NEMA-CT, and OASIS-MRI databases were employed in this study, and the findings show that the retrieval methods are effective in locating relevant information. There were three different types of biological image retrieval used: the ARP, ARR, F-Score, and ANMRR methods. It is possible to use the LBDP feature description to perform invariant face recognition.

Z Zeng [2] introduced a Novel Local Structure Descriptor (LSD) for colour picture retrieval, and local structures are constructed based on a similarity of edge orientation. No huge datasets are required for training the suggested feature extraction technique. A comparison of picture retrieval performance using the Corel database shows that descriptor to be superior. They predicted that BOW, VLAD, and compressed Fisher kernel descriptors would surpass the method in terms of performance.

Image indexing and retrieval may be improved by a set of descriptors suggested by S. A. Chatzichristofis *et al.* [8]. Evaluation of suggested descriptors is done using the Average Normalized Modified Retrieval Ranking (ANMRR). An improved retrieval score is shown when the suggested approach is applied to a big database. More reliable findings are obtained when using the descriptors FCTH and C.FCTH. If a picture has a modest number of texture regions, CEDD and C. CEDD will be able to recover it. Local diagonal extremes are found utilising first-order local diagonal derivatives to leverage the connection between diagonal neighbours of any centre pixel in a picture. Finding the values and indexes of locally occurring diagonal patterns is the goal of this method. The LDEP has been put through its paces using datasets like NEMA-CT and Emphysema-CT, and the findings show that it is both discriminative and efficient in comparison to the alternatives.

An image retrieval approach for large datasets has been suggested by J Yang *et al*. [9]. Experiments demonstrate that the approach considerably increases the sufficient performance of feature extraction and also enhances search matching. The results demonstrate that the approach has a considerable benefit in retrieval time, which may match the demands of real-time retrieval. SIFT (scale-invariant feature transform) and LBP are two novel approaches presented by L Kabbai et al. [10] for the extraction of invariant features from an interesting area (Local Binary Pattern). Different assessment metrics, such as repeatability, recall, and accuracy, are computed by these experts. It has been shown in these experiments that their approaches are more resilient than the SIFT approach in terms of repeatability, recall, and accuracy while performing different transformations such as blurring, affine transformation, and rotation. NL, NCS, SNL, SNCS, and N are some of the novel descriptors that have been tested on two datasets (LCS).

Yan Ke1 and R. Sukthankar [11] discuss the strategies. The SIFT algorithm was discovered lately by Mikolajczyk and Schmid after a thorough investigation of numerous methods. PCA-based local descriptors are more distinct and resilient against picture deformations, according to their tests. Because of its uniqueness and compactness as compared to the usual representation, PCA-SIFT improves matching accuracy (and speed) in both controlled and real-world circumstances. It is suggested that a set of colour and texture attributes would be employed for the fusion techniques, according to An Ahmed and S Mohamed [4]. Corel 1K and GHM-10K datasets were used to test the suggested technique, and the results demonstrate that the proposed method has a decent performance.

The Local Binary Pattern Histogram Fourier features (LBP-HF) presented by T. Ahonen *et al*. [12] is a new rotation-invariant image descriptor generated from discrete Fourier transforms of LBP histograms. In texture classification, material categorization, and face recognition tests, these properties surpass non-invariant and prior versions of rotation-invariant LBP and MR8. When it comes to the development of deep learning models for Content-Based Image Retrieval, S R Dubey [3] made a suggestion (CBIR). Various networks, different descriptors, and different retrieval methods are all included in their research. It is claimed that newly created models for generative adversarial networks, autoencoder and reinforcement networks as well as transfer learning have all been made available. Image retrieval using deep learning algorithms may benefit from this article.

Medical image retrieval systems developed by N F Haq *et al.* [13] have been offered for retrieving comparable pictures from a large-scale image collection. A network community identification method and a deep learning-based picture feature creation strategy are both included in the system. With an accuracy of 85%, the approach suggested was able to extract photos with identical illness labels. Content-based CT image retrieval may benefit from a novel image feature description based on the Local Wavelet Pattern (LWP) developed by S R Dubey [14]. It was evaluated on three CT image databases in terms of accuracy and recall, and the results were positive. The findings show that the suggested technique outperforms previous methods for CT image retrieval and the proposed feature descriptor's temporal complexity is smaller and more efficient and discriminative.

To attain high responsiveness, Zhou Bing and Yang Xin-xin [15] provide content-based parallel picture retrieval systems built on cluster architectures. As a way to improve the accuracy of picture matching and distortion operations like rotation and flip, their approach makes use of SCSF (Symmetric Color-Spatial Features). Users will be able to choose whether or not to utilise the texture and form features, which will be an additional option. In their paper Low-level feature extraction for visual appearance comparison between genetically changed plants for gene expression investigations, I. Yahiaoui and N. Boujemaa [16] offer a technique for botanists to analyse plant genetic changes and present it. This approach's success is shown by the researchers' findings.

A strategy termed forward analogy relevance feedback was proposed by Hui Xie *et al* [17] (Analogy-RF). An image's semantic gap between high-level and low-level characteristics may be determined by recalculating the pictures' similarity score. Using photos from the COREL-1K dataset, they were able to demonstrate the usefulness of their suggested approach. Pravin K Vikhar [18] proposed the MPEG-7, Edge Histogram Descriptor (EDH), and SVM classifier algorithms, respectively. A combination of EDH, Color auto-correlogram, colour moment, and Gabor wavelet transform is utilised to extract the pictures' information. SVM is utilised to categorise their suggested model more accurately, and the findings reveal more efficacy for their model as a whole.

M. D. Vesiü *et al.* [19] suggested a way to find images based on their content and give relevant feedback. Low-level attributes are used to characterise each picture in the database, which is based on singularity strength. It is possible to compare the performance of direct and refined feature vectors in the search. The query search may benefit from feature vectors. According to M Marinov [20], who discusses JPEG compression, various levels of compression provide diverse outcomes. It displays the findings based on a variety of similarities among the same group of images. As a consequence, the findings of two separate experiments, one including compressed and the other with non-compressed photos, indicate that the compressed pictures are more successful. H B Kekre *et al.* [21] provide a novel way for building semantic gap methods. Low-level descriptors are used mostly for content-based retrievals.

It is much more difficult to train high-level agents because they use Modified Block Truncation Coding (MBTC) to retrieve colour information and extract characteristics created by transform patterns. Feature vectors are generated as a result of the aforementioned models, which are referred to as kekre transforms. Multiple descriptors used by Luiz G *et al*. [22] provide a similarity tree-based representation with neighbours connecting each other. The quality of the visualisation of the query result has been shown by the test. Users may improve their comprehension of the CBIR system by using visualisation to better see comparable relationships between pictures, hence enabling them more control over their outcomes.

Local binary patterns are used for texture analysis and compared to the average RGB colour picture descriptor approach by Kamlesh Kumar *et al* [23]. Experiments in MATLAB reveal that precision, recall, F-measures, and retrieval time are the most important factors in determining the maximum level of accuracy when compared to a simple extraction of colours and textures. Using picture shape characteristics, c1ustering methods, and machine learning technologies, retrieval systems may become even more accurate and efficient in the future. Meenakshi M *et al*. [24] discuss the various CBIR classification and matching strategies, as well as the combination of the two. Their findings suggest that the DWT approach is more accurate, and the Canberra distance outperforms the competition. Complexity of time and effective model design are the two biggest challenges in CBIR.

S.S. Reddy *et al* [25] worked on Object detection by using various machine learning algorithms like YOLO [26], Medical images security in Cloud [27], removing denoising by using PSO and APSO [28], recognition by using flower species [29], Image Coloriozation using CNN [30], classification of gender by voice [31] and by using HMM [32], sharpening Blur Image [33], Facial expression by using Beizer Curves [34], image retrieval using XML search [35], retrieving an image using GA [36], subsequence identification on Videos [37].

Using the pseudo-scanner standardisation (DI-PSS) methodology, Hayato Arai1 *et al*. [38] developed a disease-oriented image embedding method. Data harmonisation and an algorithm known as 3D convolutional autoencoders are the two main models in this framework (3D-CAE). Scans and datasets are responsible for 15.8-22.6 percent of the differences between Alzheimer's patients and those with clinically normal brains. VGG19-based novel feature extraction was used by Z. Nawab Khan et al. [39] in their CBIR system is completely automated. Using transfer learning, they also proposed a way to improve the retrieval performance from the query and dataset images. For their experiments, they utilise the CE-MRI dataset. An average accuracy of 96.13% may be expected. On TI-weighted CE-MRI, CBIR is carried out. Comparing TI-weighted and non-TI weighted photos, the findings are positive. Precision, memory, and accuracy may all benefit from fine-tuning.

For multi-label RS CBIR challenges, G Sumbul *et al*. [40] suggested a new triplet sampling strategy in the context of deep neural networks designed for that purpose. Two steps are required to complete the procedure. The anchors are a necessary part of the initial setup. By analysing the hardness, and the efficiency of the pictures in both the query image and dataset images, the second algorithm has distinct placements of positive and negative images in the dataset. Experiments reveal that Deep neural networks outperform simpler networks in terms of performance. Measures like reproducibility and recall are some of the metrics they use to assess an experiment.

3. PROBLEM STATEMENT

All existing techniques send image queries in plain text format, and the cloud may monitor such data and can misuse that data. So to avoid such misuse image and query encryption algorithms were introduced. Still, these algorithms are insufficient as this algorithm requires decryption before performing a content-based image search, and this decryption can leak data. In propose SeCBIR (secure content-based image retrieval) using the VGG16 neural network algorithm to extract features rather than directly using images. If we used explicit images, the cloud might misuse this information, so we are training VGG16 with the COREL10K images dataset. VGG16 model can be used to extract features from any QUERY image and then apply PAHE encryption to secure components. Then these encrypted features will be sent to the cloud, and the cloud will directly perform specific content-based images.

4. METHODOLOGY

The details of objectives, dataset and system architecture of the proposed methodology were described in this section.

a. Dataset Description

The dataset contains 900 different images. 720 images are used for training, and the remaining images are used for testing. The image set consists of 10 different folders containing 100 other images. In data pre-processing, it will read all images, normalise image pixel features, remove noise from the dataset, and then identify the missing values in the dataset. The images of the dataset are shown in Figure-4.

VOL. 18, NO. 3, FEBRUARY 2023 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

Figure-4. Images of dataset.

b. Objective

The main objectives of the work are to find the most relevant image retrieval from the given dataset using an effective algorithm to calculate its accuracy.

c. System Architecture

Figure-5 represents the proposed model's system architecture. Firstly, the image dataset is loaded, and then image pre-processing is done.

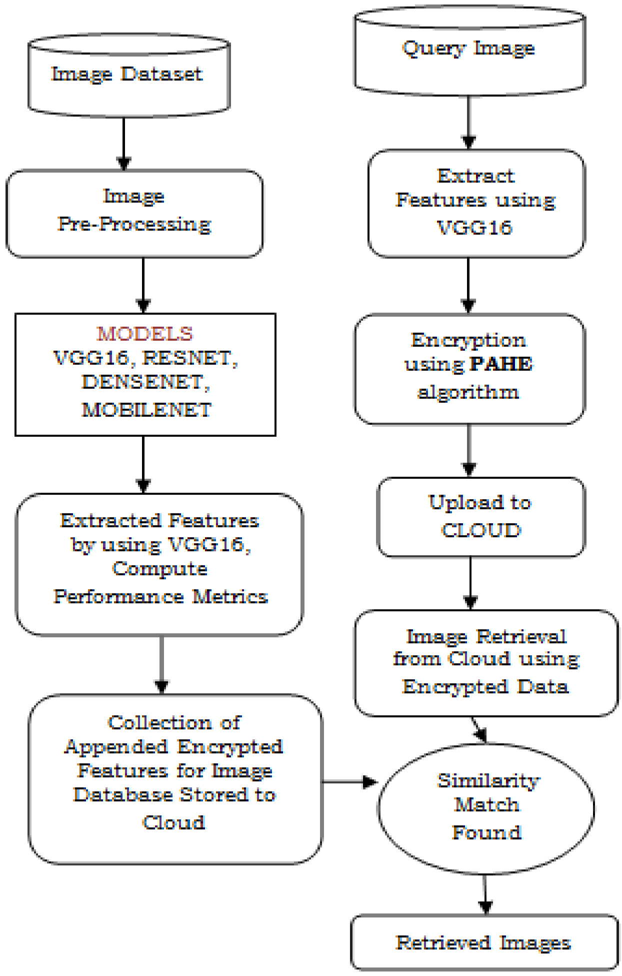

Figure-5. System architecture.

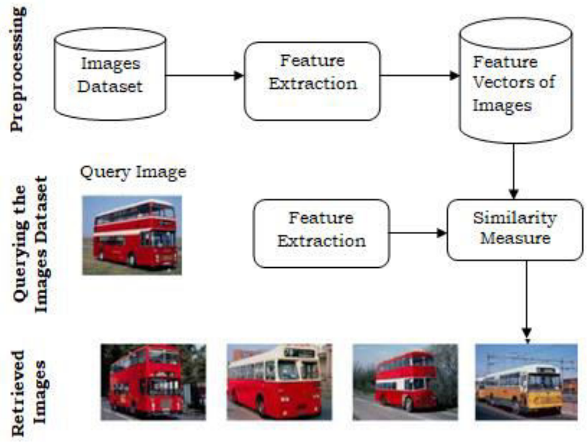

Figure-6. Content-based image retrieval.

For training, we have taken 80% of the data and 20% for testing and then used the VGG16 neural network algorithm to extract features rather than directly using images. If we used explicit images, the cloud might misuse this information, so we are training VGG16 with the COREL10K images dataset. Figure-6 shows the content-based image retrieval system architecture.

VGG16 model can be used to extract features from any QUERY image and then apply PAHE encryption to secure components. Then these encrypted features will be sent to the cloud, and the cloud will directly perform certain content based on image similarity on encrypted characteristics to perform an image search and then send the search result back to the user/client.

d. Algorithms Used

RESNET

Deep networks are hard to train because they have vanishing gradient problems. Residual networks are deep networks, but we will not face vanishing gradients while using the residual networks. When preparing the images, we may face issues, so residual nets have the concept of skip connections, which means adding the original or actual input to the output of the convolutional block means it wants to make inputs and outputs the same. It has to start with a convolutional layer of filter size 7*7, the stride of 2 and padding size is three, and next having max pooling layer of filter size 3*3, stride 2 and padding remains same, and the formula for both the layers is (n+2p-f/s) +1. From the diagram, the dotted lines represent the skip connections. ResNet has different types of variants there are ResNet18, ResNet50, ResNet101, ResNet56, etc. Here they have the same concept but with several different layers. After all the connected blocks, it has a fully connected layer with a value of 1000. Figure-7 shows the architecture of ResNet.

ARPN Journal of Engineering and Applied Sciences

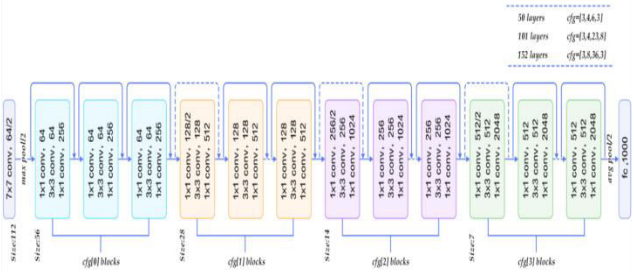

Figure-7. ResNet architecture.

DENSENET

DenseNet (Densely connected convolutional networks) also uses the same concept of convolutional neural networks with the same layers of convolutional, pooling, and RELU activation functions, and the architecture of Densenet is shown in Figure-8. Each layer takes the features as inputs from the previous layer, which helps the training process to remove the vanishing gradient problems. These vanishing gradient problems are present in every deep neural network. When we backpropagate the errors, every step reduces these errors. The critical feature of this architecture is the dense blocks. These dense blocks are inter-connected, which is helpful for feature reuse and reducing the number of parameters because it uses the previous feature map information instead of generating more parameters. It reduces the overfit, and this algorithm works more accurately than ResNet.

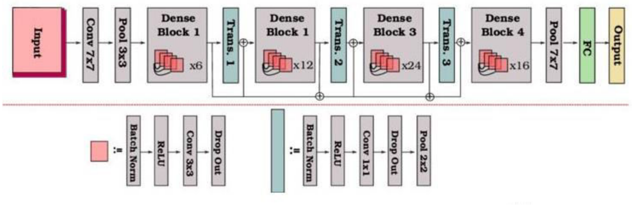

Figure-8. Densenet architecture.

MOBILENET

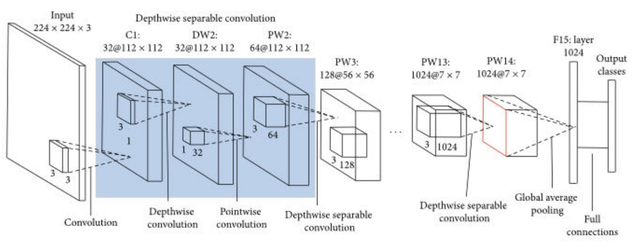

Figure-9. Densenet architecture.

MobileNet follows two types of convolutional functions one is depth-wise, and the other one is point-wise separable convolutions, and the architecture of Densenet is shown in Figure-9. Depth-wise convolutional function works as shown in Figure-10.

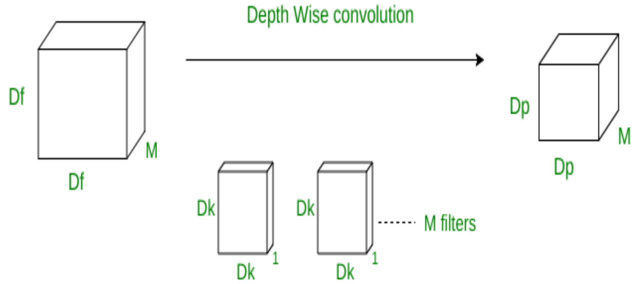

Figure-10. Depth-wise convolution.

In-depth, f and k represent the filters, kernels are the input size of Dp*Dp*1, and the convolution is applied to every channel. Then the output will be the size of Dp*Dp*M.

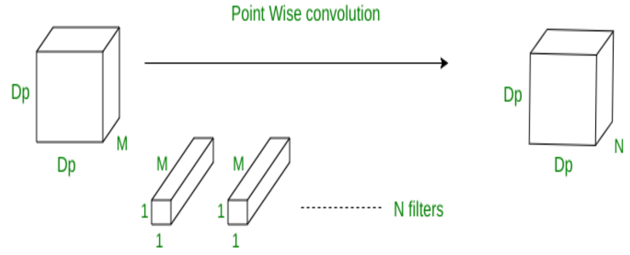

Figure-11. Point-wise convolution.

In point-wise function, a 1*1 convolution operation is applied to the M channels, then the filter size for the process is 1*1*M, and the output size becomes Dp*Dp*N. The point-wise convolutional function works as shown in Figure-11.

e. Proposed Algorithm VGG16

The vgg16 architecture consists of 4 layers, as shown in Figure-12. They are the convolutional layer, max pooling layer, fully connected layer, and softmax layer. When the image is getting passed through the convolutional layer, the convolutional layer count is 2 because there are two convolutional layers and then one max pooling layer. Again, it has 2 convolutional layers, and 1 max pooling layer and so on. It has 2-1-2-1-3-1-3-1-3-1-fc-fc-fc-softmax. Fc represents the fully connected layer and, lastly softmax layer. Here the image sizes 224*224*3 and 224*224 represent the image's length and breadth, and 3 illustrates RGB.

ARPN Journal of Engineering and Applied Sciences

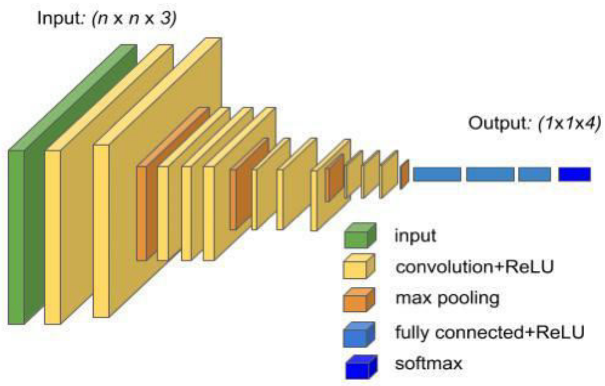

Figure-12. VGG16 architecture.

In vgg16, it has fixed values of filter size 3*3, the stride of 1 and padding remains same, and coming to max pooling, it has a filter size of 2*2, stride 2 and padding is same. To calculate the values in the max pooling layer, it has a formula of (n+2p-f/s)+1, and in the convolutional layer, dimensions remain the same, but filter sizes are applied. Here 3*3 represents the 64 pixels. After that, it has 3 fully connected layers and one softmax layer.

Proposed Algorithm-1 for VGG16	
Step 1:	Loading the VGG model from the saved file.
Step 2:	Loading VGG model weights and assigning to classifier variable.
Step 3:	Make VGG classifier ready for prediction.
Step 4:	Extracting the last layer from the VGG classifier.
Step 5:	Using the VGG classifier to extract its input and output and assign them to the model variable.
Step 6:	Now the VGG model is ready to extract features using the VGG model object.
Step 7:	Now encrypt those features by using the PAHE algorithm.
Step 8:	Next, start looping all the images.
Step 9:	Read the image from each directory and assign the image value to the Input variable.
Step 10:	Calling predict function from the VGG model to extract features from the input image.
Step 11:	Now predicted features will get encrypted and then used this features for image retrieval
Step 12:	Save the features.

Firstly loading the VGG model from the saved files and then loading weights assigned to the classifier. Start implementing the classifier by extracting the last layer from the model to begin predicting consequences for each layer by giving input o know the output. Now it is ready to read all images from the directory and provides image values to the input variable by calling the predict function from the VGG model to extract features from the input image. Now saving the features for encrypting data means encrypting data. It takes inputs as extracted features from the model and starts looping all the images, protecting them, encrypting the data, and the process repeats.

Algorithm Steps for PAHE:

Homomorphic encryption can perform image searches directly on encrypted data without decryption, but this algorithm is heavy in computation.

So we are using a lightweight PAHE scheme, which can directly speed up linear algebra and has shown its effectiveness.

Polynomial Function for Encrypting Data

$M1, m2 \in Z$ then encryption function of

$e(m1)+e(m2)=e(m1+m2)$ & $e(m1)e(m2)=e(m1m2)$

First, two cypher texts are encryptions of the same value if and only if they lie in the same coset of $(f, g) \subset Z[x, y]$.

The encryptions $e(m1)$ and $e(m2)$ are of the form

$$e(m1) = m1 + a1f + b1g$$

$$e(m2) = m2 + a2f + b2g$$

for some $ai, bi \in Z[x, y]$.

Then $e(m1) + e(m2) = m1 + m2 + (a1 + a2)f + (b1 + b2)g$

This is in the coset $m1 + m2$ of (f, g), so it is encryption of $m1 + m2$.

This proves that the scheme is additively homomorphic.

Proposed Algorithm-2 for PAHE	
Step 1:	Taking inputs as extracted features from VGG.
Step 2:	Then applying a mathematical polynomial function to image pixels to encrypt data.
Step 3:	Start the process of generating a random key.
Step 4:	Wrote part to encrypt data.
Step 5:	Then decrypting data with a private key.
Step 6:	Encoding the integer into a plaintext polynomial.
Step 7:	For decryption, make a reverse polynomial function on features to decrypt the image.

ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

5. RESULTS

Table-1. Evaluation of Various Metrics.

S.No	Metrics	Equation/Formulae
1	Accuracy	$Ay = \frac{(TP+TN)}{(TP+TN+FP+FN)}$
2	Precision	$Pn = \frac{(TP)}{(TP+FP)}$
3	Recall	$Rl = \frac{(TP)}{(TP+FN)}$
4	F-Score	$Fe = \frac{(TP)}{(TP+\frac{1}{2(FP+FN)})}$
5	Jaccard Index	$J(A,B) = \frac{\lvert A\cap B\rvert}{\lvert A\cup B\rvert}$
6	Error Rate	$Ert = \frac{(FP+FN)}{(P+N)}$
7	AUC-ROC Curve	$AR\ crv = \frac{(1+TPR-FPR)}{2}$

From Table-1, various metrics were evaluated using the Confusion matrix of multiple models to know the performance of the Models.

Results Obtained for VGG16

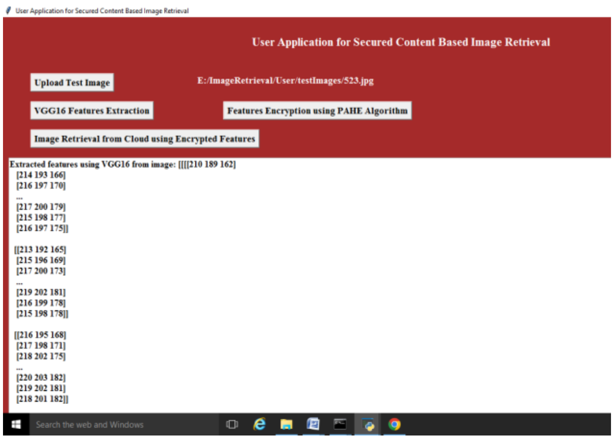

Figure-13. Screenshot of extracted pixel values.

In Figure-13, image pixel values are extracted from VGG16. Now click on the 'Features Encryption using PAHE Algorithm' button to encrypt features and get the below output.

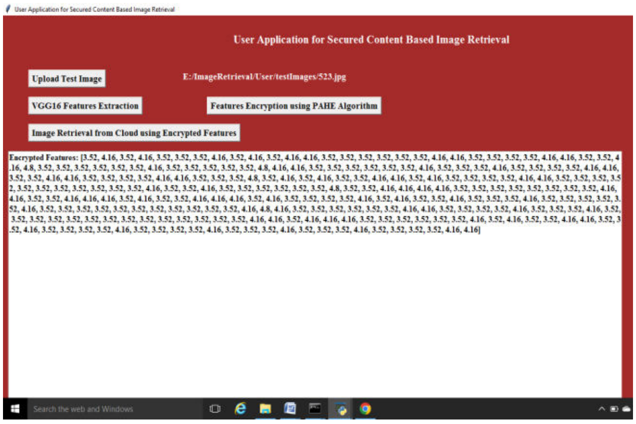

Figure-14. Converted integer encrypted values from features.

Figure-14 shows that image features are converted to encrypted integer values. Now click the 'Image Retrieval from Cloud using Encrypted Features' button to send encrypted parts to the cloud and get the search result.

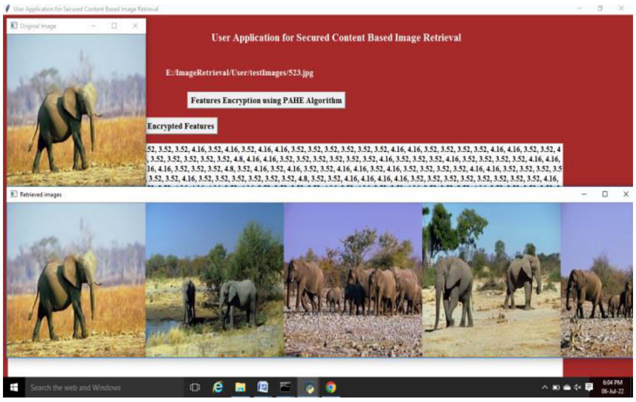

Figure-15. Query and Content Images.

In Figure-15, the top image is the query image, and the below images are similar content images returned by the cloud. Similarly, you can upload other images, extract features, encrypt features, and then upload them to the cloud to get the search result back.

Table-2. Obtained results from various metric from various models.

Model Name	Performance Metrics					
	Accuracy	Precision	Recall	F-Score	Error rate	Jaccard Index
RESNET	70.55555	61.05555	71.0243	69.90954	19.44444	79.419753
VGG16	99.4444	99.49999	99.4999	99.4871	0.55555	98.91666
DENSENET	89.44444	91.68605	87.4198	87.8355	10.55555	81.52128
MOBILENET	76.1904	83.943	77.287	73.333	3.937471	86.6666

ARPN Journal of Engineering and Applied Sciences

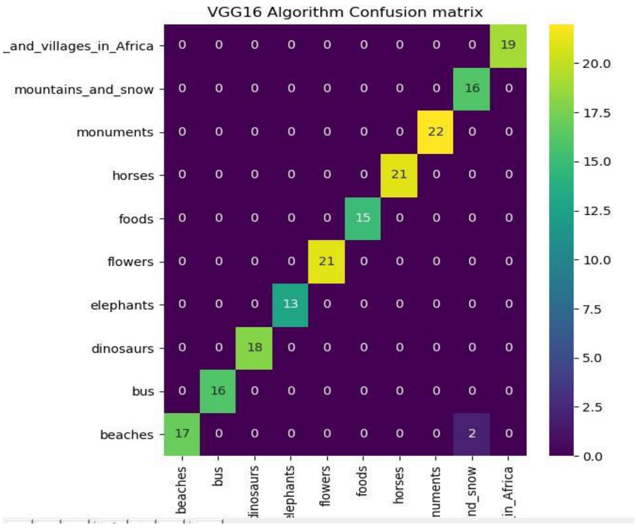

Figure-16. Confusion matrix obtained to VGG16.

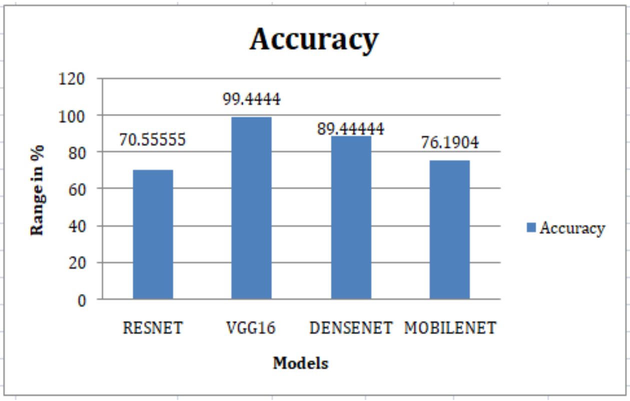

Figure-17. Accuracy for various models.

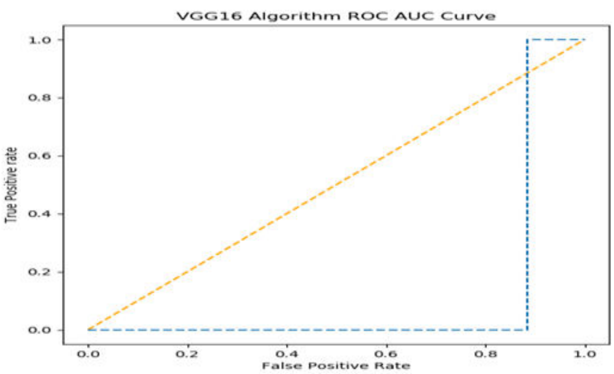

Figure-18. AUC and ROC curve for VGG16.

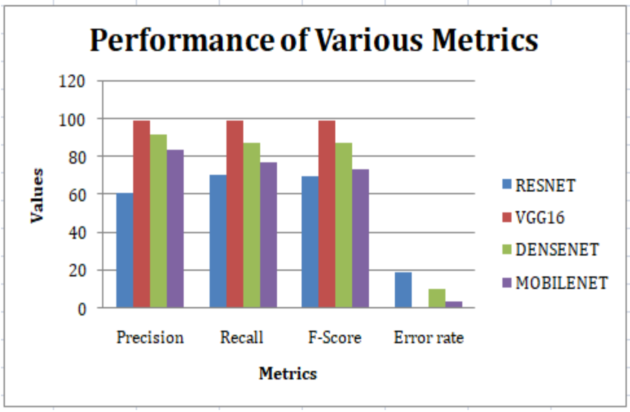

Figure-19. Performance analysis of various metrics.

Figure-16 is a Confusion Matrix of VGG16. By using this Matrix, Table-2 was generated on various Metrics by using Table-1. Comparison of Accuracy with Various Models like RESNET, DENSENET, VGG16, and MOBILENET, shown in Figure-17. Among them, 99.44% was the highest accuracy acquired for VGG16. AUC & ROC curve was drawn for VGG16 as shown in Figure 18, and Various Performance evaluation metrics were evaluated on various Models like Precision, Recall, F-score, and Error rate. Among them, the proposed VGG16 got the highest %, shown in Figure-19.

6. CONCLUSIONS

In this paper, we proposed a CBIR process using the COREL-10K dataset consisting of 10 different folders. We have implemented it here using 4 other models: DENSENET, RESNET, MOBILENET, and VGG16. Select the best model by calculating evaluation metrics for each algorithm. Those metrics are accuracy, precision, recall, F1-score, Error Rates, Jaccard index and the values are *99.4444, 99.49999, 99.4999, 99.4871, 0.55555,* and *98.91666* respectively. We can say that VGG got the best model based on these values. So, VGG16 is used to extract features from the image dataset, and PAHE is used to encrypt attributes. Then secure image computations will be performed on the cloud side to compare the two images. Then the cloud will directly perform certain content-based image similarities on encrypted features to perform an image search and then send the search result back to the user/client.

REFERENCES

[1] Liu G. H., Li Z. Y., Zhang L, Xu Y. 2011. Image retrieval based on micro-structure descriptor. Pattern Recognition. 44(9): 2123-33.

[2] Zeng Z. 2016. A novel local structure descriptor for color image retrieval. Information. 7(1): 9.

[3] Dubey S. R. 2021. A decade survey of content based image retrieval using deep learning. IEEE

Transactions on Circuits and Systems for Video Technology. 32(5): 2687-704.

[4] Ahmed A., Mohamed S. 2021. Implementation of early and late fusion methods for content-based image retrieval. International Journal of Advanced and Applied Sciences. 8(7): 97-105.

[5] Murala S., Maheshwari R. P., Balasubramanian R. 2012. Local tetra patterns: a new feature descriptor for content-based image retrieval. IEEE transactions on image processing. 21(5): 2874-86.

[6] Zhao M., Zhang H., Sun J. 2016. A novel image retrieval method based on multi-trend structure descriptor. Journal of Visual Communication and Image Representation. 38: 73-81.

[7] Dubey S. R., Singh S. K., Singh R. K. 2015. Local diagonal extrema pattern: a new and efficient feature descriptor for CT image retrieval. IEEE Signal Processing Letters. 22(9): 1215-9.

[8] Chatzichristofis S. A., Zagoris K., Boutalis Y. S., Papamarkos N. 2010. Accurate image retrieval based on compact composite descriptors and relevance feedback information. International Journal of Pattern Recognition and Artificial Intelligence. 24(02): 207-44.

[9] Yang J., Jiang B., Li B., Tian K., Lv Z. 2017. A fast image retrieval method designed for network big data. IEEE Transactions on Industrial Informatics. 13(5): 2350-9.

[10] Kabbai L., Azaza A., Abdellaoui M., Douik A. 2015. Image matching based on LBP and SIFT descriptor. In2015 IEEE 12th International Multi-Conference on Systems, Signals & Devices (SSD15) 16 (pp. 1-6). IEEE.

[11] Ke Y., Sukthankar R. 2004. PCA-SIFT: A more distinctive representation for local image descriptors. In Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR (Vol. 2, pp. II-II). IEEE.

[12] Ahonen T., Matas J., He C., Pietikäinen M. 2009. Rotation invariant image description with local binary pattern histogram fourier features. In Scandinavian conference on image analysis 2009 Jun 15 (pp. 61-70). Springer, Berlin, Heidelberg.

[13] Haq N. F., Moradi M., Wang Z. J. 2021. A deep community based approach for large scale content based X-ray image retrieval. Medical Image Analysis. 68: 101847.

[14] Dubey S. R., Singh S. K., Singh R. K. 2015. Local wavelet pattern: a new feature descriptor for image retrieval in medical CT databases. IEEE Transactions on Image Processing. 24(12): 5892-903.

[15] Bing Z., Xin-xin Y. 2010. A content-based parallel image retrieval system. In 2010 International Conference on Computer Design and Applications (Vol. 1, pp. V1-332). IEEE.

[16] Yahiaoui I., Boujemaa N. 2005. Content-based image retrieval in botanical collections for gene expression studies. In IEEE International Conference on Image Processing 2005(Vol. 3, pp. III-1240). IEEE.

[17] Xie H., Ji Y., Lu Y. 2013. An analogy-relevance feedback CBIR method using multiple features. In2013 International Conference on Computational Problem-Solving (ICCP) (pp. 83-86). IEEE.

[18] Vikhar, Pradnya and Pravin Karde. 2016. Improved CBIR system using edge histogram descriptor (EHD) and support vector machine (SVM). In 2016 International Conference on ICT in Business Industry & Government (ICTBIG), pp. 1-5. IEEE.

[19] Vesić M. D., Zajić G. J., Gavrovska A. M. 2016. CBIR system with relevance feedback using Hölder exponent. In 2016 13th Symposium on Neural Networks and Applications (NEUREL) (pp. 1-4). IEEE.

[20] Marinov M. 2020. Comparative Analysis on Different Degrees of JPEG Compression Used in CBIR Systems. In 2020 XI National Conference with International Participation (ELECTRONICA) (pp. 1-4). IEEE.

[21] Kekre H. B., Bharadi V. A., Sawant A. R., Bijith M. 2012. Fusion of Modified Block Truncation Coding and Kekre transform patterns for colour & texture based CBIR. In 2012 International Conference on Communication, Information & Computing Technology (ICCICT) (pp. 1-6). IEEE.

[22] Real L. G., Bueno R., Ribeiro M. X. 2019. Evaluating boundary conditions and hierarchical visualization in cbir. In 2019 23rd International Conference Information Visualisation (IV) (pp. 68-73). IEEE.

[23] Kumar K., Li J. P. 2015. Complementary feature extraction approach in CBIR. In 2015 12th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP) (pp. 192-197). IEEE.

[24] Madugunki M., Bormane D. S., Bhadoria S., Dethe C. G. 2011. Comparison of different CBIR techniques. In2011 3rd International Conference on Electronics Computer Technology (Vol. 4, pp. 372-375). IEEE.

[25] Shankar R. S., Gupta V. M., Murthy K. V., Someswararao C. 2012. Object oriented fuzzy filter for noise reduction of Pgm images. In2012 8th International Conference on Information Science and Digital Content Technology (ICIDT2012) (Vol. 3, pp. 776-782). IEEE.

[26] Shankar R. S., Srinivas L. V., Neelima P., Mahesh G. 2022. A Framework to Enhance Object Detection Performance by using YOLO Algorithm. In2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS) (pp. 1591-1600). IEEE.

[27] Karunasri K., Mahesh G., Shankar R. S. 2022. Medical images security in cloud computing using cp-abe algorithm. ARPN Journal of Engineering and Applied Sciences. 17(7): 759-766.

[28] Gupta V. M., Murthy K. V., Shankar R. S. 2021. A Novel Approach for Image Denoising and Performance Analysis using SGO and APSO. InJournal of Physics: Conference Series (2070(1): 012139). IOP Publishing.

[29] Shankar R. S., Srinivas L. V., Raju V. S., Murthy K. V. 2021. A Comprehensive Analysis of Deep Learning Techniques for Recognition of Flower Species. In 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV) (pp. 1172-1179). IEEE.

[30] Shankar R. S., Mahesh G., Murthy K. V., Ravibabu D. 2020. A Novel approach for Gray Scale Image Colorization using Convolutional Neural Networks. In 2020 International Conference on System, Computation, Automation and Networking (ICSCAN) (pp. 1-8). IEEE.

[31] Shankar S., Raghaveni J., Rudraraju P., Vineela Sravya Y. 2020. Classification of gender by voice recognition using machine learning algorithms. Journal of Critical Reviews. 7(9): 1217-29.

[32] Vmnssvkr Gupta, R. Shiva Shankar, Harika Devi Kotha, J. Raghaveni. 2020. Voice Identification in Python Using Hidden Markov Model. International Journal of Advanced Science and Technology. 29(06): 8100-8112

[33] Shankar R. S., Mahesh G., Murthy K. V., Rajanikanth J. 2020. A novel approach for sharpening blur image using convolutional neural networks. Journal of Critical Reviews. 7(7): 139-48.

[34] Babu D. R., Shankar R. S., Mahesh G., Murthy K. V. 2017. Facial expression recognition using bezier curves with hausdorff distance. In2017 International Conference on IoT and Application (ICIOT) (pp. 1-8). IEEE.

[35] Rajanikanth J., Shankar R. S., Raju V. S., Rao C. S. 2016. A Novel Mechanism for Information Retrieval with XML Keyword Search. I J C T A. 9(36): 685-692.

[36] Shankar R. S., Sravani K., Srinivas L. V., Babu D. R. 2017. An approach for retrieving an image using Genetic Algorithm. International Journal of Latest Trends in Engineering and Technology. 9(8): 057-64.

[37] Shankar R. S., Krishna A. B., Rajanikanth J., Rao C. S. 2012. Implementation of object oriented approach to query processing for video subsequence identification. In2012 National Conference on Computing and Communication Systems (pp. 1-5). IEEE.

[38] Arai H., Onga Y., Ikuta K., Chayama Y., Iyatomi H., Oishi K. 2021. Disease-oriented image embedding with pseudo-scanner standardization for content-based image retrieval on 3D brain MRI. IEEE Access. 9: 165326-40.

[39] Swati Z. N., Zhao Q., Kabir M., Ali F., Ali Z., Ahmed S., Lu J. 2019. Content-based brain tumor retrieval for MR images using transfer learning. IEEE Access. 7: 17809-22.

[40] Sumbul G., Ravanbakhsh M., Demir B. 2022. Informative and Representative Triplet Selection for Multi-Label Remote Sensing Image Retrieval. IEEE Transactions on Geoscience and Remote Sensing.