WJCMS

Wasit Journal of Computer and Mathematics Science

Journal Homepage: https://wjcm.uowasit.edu.iq/index.php/WJCM

e-ISSN: 2788-5879 p-ISSN: 2788-5879

Classification on Unsupervised Deep Hashing With Pseudo Labels Using Support Vector Machine for Scalable Image Retrieval

Rohit Sharma[1,*], Bipin Kumar Rai[2] and Shubham Sharma[3]

[1]Department of Electronics & Communication Engineering, SRM Institute of Science and Technology, NCR Campus, Ghaziabad, UP, India
[2]ABES Institute of Technology, Ghaziabad, India
[3]Mechanical Engineering Department, University Center for Research & Development, Chandigarh University, Mohali, Punjab, India

*Corresponding Author: Rohit Sharma

DOI: https://doi.org/10.31185/wjcm.147
Received: March 2023; Accepted: June 2023; Available online: June 2023

ABSTRACT: Users can formulate their queries more easily using content-based image retrieval (CBIR). Still, it also produces poor retrieval results due to its focus on the visual attributes of user-input query objects. CBIR works on automatically signing keywords with images, which supports image retrieval, and image annotation was suggested as the best framework. Using an annotated image dataset, this paper aims to solve the problem of CBIR using deep learning techniques, particularly Convolutional Neural Networks (CNNs).

Keywords: CNN, Deep Learning, Image Retrieval, SVM.

1. INTRODUCTION

The development of multimedia & computer tools has led to the manufacturing of computerized images & low-priced, huge image sources. The size of the image collections has been improved quickly, comprising digital collections, therapeutic images, etc. To switch this quick progress, developing image retrieval schemes that work on an enormous measure is mandatory. The essential point is constructing a strong system that makes, oversees, and inquires about image catalogues in an exact path. CBIR is a technique of consequently collecting images by the abstraction of the low-level graphics highlights, comparable to shape, surface and color as well as these recorded highlights completely in the responsibility of retrieving images [1], [2]. In this manner, it very well may be said that through route, perusing, and inquiry by-precedent and the closeness among the low-level image constituents, that can be employed to recover significant images. The images symbolize the points in the high-dimensional constituent space measurement utilized to compare or contrast images in the space.

Quick comparability checking is one of the key prerequisites for large-scale visual data recovery applications [3–7]. Hashing is the most commonly applied technique for high-dimensional space element graphics to uncover a low-dimensional space-saving resemblance. The muddle codes greatly minimize the need for extra room to improve mathematical abilities [8], [9]. As of late, many hashing methods [10–13] have carried out impressive executions.

The recovery of a material-based object recovery system is essentially based on an element representation and similarity approximation that has been widely viewed by visual and sound scientists for a considerable period. Even though various programs have been introduced, the CBIR visits stick out between the more research problems and the CBIR flows inquiry. From a unique point of view, such a check may be added to the critical AI experiment, that is, how knowledgeable computers, such as humans, are installed or prepared to handle approved tasks [10].

The CNN mentions classes, where various layers of in-arrangement handling stages in various levelled models are demoralized for pattern classification & feature illustration learning [1], [10]. Convergence of several areas of neural

*Corresponding author: rohit.453.sh@gmail.com

systems research, graphic presentation, development, model recognition, signal processing, and so on [11] earned Deep controlled back-propagation, CNN arranged for digit recognition.

It has become a lucrative work concept in PC vision ML, where Deep learning produces outcomes for several tasks. CNN also recommended that the ImageNet Arrangement with CNN be shown in the image classification undertaking [11]. The template planned for more than one million images, producing five task blunders within 15.3 % of more than 1,000 students. Overall, some of the new tasks showed improvement in the CNN method. The training system's key five error rate decreased to 13.24 % to immediately categorize, identify and differentiate items. Apart from grouping an image, an object recognition errand can also benefit from the CNN model, as stated in more detail. As a rule, three main objectives behind the renown of deep adaptation today are the dramatically expanded chip handling efficiency, the profoundly low price of hard-product registration, and the late development of AI sign/data work. Recently, CNN methods have been suggested and broadly examined, DBN, BM, RBM, DBM, and DNN [12]– [14]. Among different systems, the Deep CNN, which is Deep discriminative engineering and has a place with the DNN class, has cutting-edge execution on the different assignments and rivalries in PC vision and image acknowledgement.

CNN's approach for resolving the CBIR undertaking for images which have been explained. There was no investigation of the methodology being connected to datasets. In this study, three types of datasets are used such as WANG, Flower, and CIFAR. Deep learning to apply with best class Deep learning strategy, that knows CNN for taking in highlight portrayals from image information. Here organize a large-scale deep convolutional neural system for learning successful element portrayals of images. Each image will be relegated as a double incentive for every class present in it. Every image allotted with the 8-bit double number where every piece speaks to every class stored in a list. When an inquiry image is given, it will likewise be prepared and 8-bit binary number for it will be created and coordinated against the file. Every one of the images with the coordinating twofold returned as the nearest coordinate for the query image

2. RELATED WORK

CBIR utilizes image content highlights to retrieve computerized images from a huge database. Various visual element extraction systems have been utilized to actualize the seeking reason. Because of the necessity of calculation time, some great calculations are not utilized. The recovery accomplishment of an ingredient-based image recovery system urgently relies upon element portrayal and comparability estimations. A definitive point of the proposed strategy is to furnish a proficient calculation to manage the previously mentioned issue definition. Here the DBN strategy for profound learning is utilized to extricate the highlights and order and is a developing exploration region because of the age of huge volume of information. The proposed technique is tried through recreation in correlation, and the outcomes demonstrate a tremendous positive deviation towards its exhibition [15–17].

Here utilized a blend of the predominant colours, normal colours, irresistible channels, and fuzzy colours histogram for portraying colours highlight. The fuzzy 3D colours are a prerequisite to figuring out prevailing colours. The fuzzy form is progressively adjusted for hues that fall among colour receptacles. We have utilized just eight colour receptacles in this task [18].

Learning-based hashing (LBH) strategies register binary code manipulating and preparing the information. Its related exclamation marks keep away from the disservices of information-free techniques, for example, LSH that translates input information so comparable information was anticipated into similar basins with high probability, and this sort of strategy was demonstrated not strong [19].

To accomplish time reduction, multifaceted nature and better appropriateness furthermore roused by ghostly hashing technique. It can consequently find area characteristic systems in the preparation information. The hyperplane-based hash work, also called Spherical Hashing, limits the round separation among the first genuine esteemed highlights. The binary code separation measurements are called circular Hamming separation [20], [21].

For the other classification, the interpretation name data of every example is completely utilized, adapting increasingly experienced binary representation; along these lines, preferable presentation over the unsupervised and semi-directed strategies can be accomplished. For instance, uses pairwise families among the information tests to limit recreation mistakes among the first Euclidean distance and mastered Hamming space [22], [23].

A sack of pictorial words [9] in computer vision is characterized as a path of event tallies of the terminology of the nearby image highlights [2], [24]. To register, the key focuses initially utilized SURF and afterwards contrasted the outcomes and SIFT that worked enhanced. SURF is a strong neighbourhood, including a locator. It utilizes a whole number estimate to the determinant of the Hessian mass indicator, which can be rapidly processed with a vital image (3 number activities), which utilizes the Hessian Threshold as 600. The filter is a calculation to identify and portray the neighbourhood included in images. The neighbourhood image inclinations are estimated at the chosen scale in the area around all the points. This change portrayal considers huge dimensions of neighbourhood shape contortion enlightenment.

3. METHODOLOGY

CNN is one feed-forward ANN where each uniquely placed neuron can reach overlapping regions in the visual field. Biologically-inspired invariants of MLP Are intended to minimize pre-processing, which models are extensively used in image and video identification. At the time of image recognition, CNN seeks tiny slices of the image given as input and termed as receptive fields in assistance to numerous layers of the small neuron collections of the model [10], [17], [25]. Convolutional networks may include the output of neuron clusters with a combination of local or total merging layers. In context with the biological process, Layers connected fully and convolutional are equipped in convolutional networks, with a point-to-point nonlinearity added at the termination or next to every layer. The convolution operation is applied on minor sections to prevent the scenario, whereas uncountable parameters will exist in a case where each layer is connected fully. Convolutional networks distribute weights in the convolutional layers, which are advantageous to reduce the memory size required and increase performance as used in the same filter layer or weights bank for each pixel. CNNs use comparatively less amount of pre-processing in comparison with other image grouping algorithms [26], [27].

It implies that in other systems, which are traditionally manually designed, the network knows about filters. Because of the lower restriction on prior knowledge, CNN has a major benefit over others, and designing hand-engineered characteristics is difficult [13], [14], [28].

Likelihood Maximization techniques accept the examples in various classes share some inactive traits to guarantee the pseudo names can genuinely reflect the circulation of the real class marks, and probability boost is utilized to amplify the joint conveyance between visual highlights and pseudo names. Quantization Error reduction methods [29] are enhanced for better quantization. Here best in class technique ITQ [20], [21], which utilized symmetrical revolution network R to extend every datum point to the closest vertex of the binary hypercube, and after that, R is refreshed to limit the quantization loss as indicated by this task.

The correlation-enhancing method ITQ-CCA [30] expects to augment the binary code and visual component relationship. 4) DH maintains a strategic distance from utilizing hand-made highlights, for example, GIST [3] and SIFT, and makes the element extraction and encoding into a brought-together start-to-finish model.

Let $X = [x_1; \ldots; x_n; \ldots; x_{N]}$ € $R^{N\times w\times h}$ be the preparation set which contains N image tests, where x_n € $R^{w\times h}$ is the n-th image test in X and has the element of w ×h. Hashing strategies expect to get familiar with a hash work H: X→B €$\{-1, 1\}^{N\times k}$ to extend each example into a minimal twofold code $b_n = H(x_n)$, where k is the objective parallel code length.

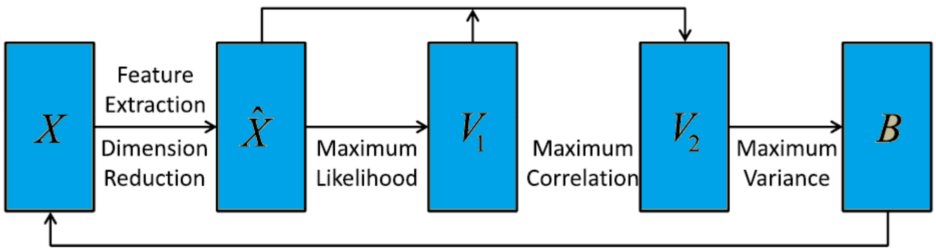

FIGURE 1. The main system of Unsupervised DH technique with Pseudo Labels.

The found pseudo-name codes are supervised for tweaking the profound hashing system like existing administered strategies. The all-out strategy is executed for T times intermittently to discover a better system model.

Figure 1 delineates the structure of deep learning hashing strategy. The essential issue is how to get class marks amid preparing. In administered strategies, the class names are utilized to direct intra-class occasions to be encoded nearer in B. For improved execution under unsupervised situations, this paper proposes finding dormant pseudo marks that can mirror the information circulation.

Correspondingly, we receive three of the most important criteria to adjust the inert names on numerous occasions. Initially, the inert names V_1 € $R^{N\times k0}$ augment the probability to the visual element ^X € $R^{N\times d}$, which is disengaged from visual example X. Be that as it may, in many situations, stable execution frequently results from fewer dormant marks, for example, $k_0 < k$, which may prompt extreme data misfortune. In this way, the subsequent stage finds sufficient inert classes V_2€ $R^{N\times k}{}_1$ that can save the most extreme shared data, where $k_1 = k$ is utilized as the length of conclusive hash codes. The latent labels are changed over into parallel codes that can safeguard the appropriation of V_2 meanwhile augmenting the general difference with the goal that the last hash codes are progressively discriminative. From this point forward, we can intermittently prepare profound hashing models utilizing the found inactive name codes like administered techniques. The complete flow chart of the existing methodology is presented in Figure 2.

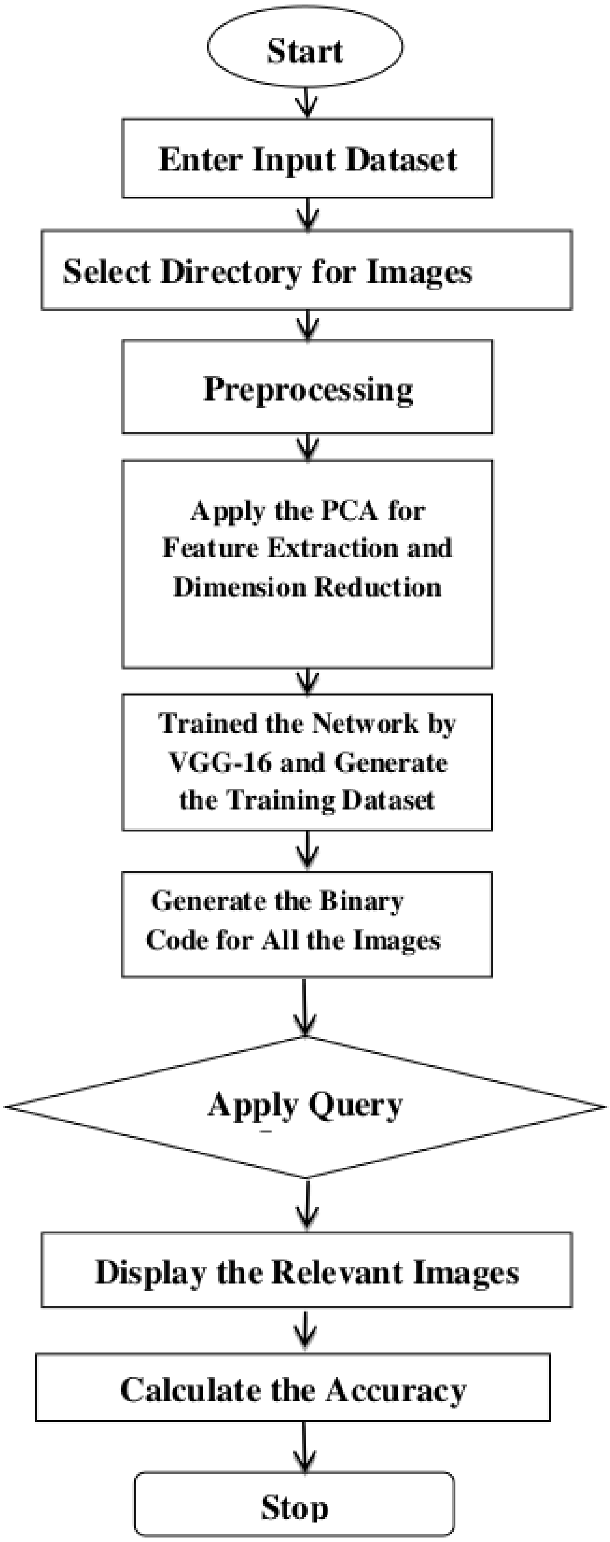

FIGURE 2. The flow chart for the deep learning hashing method.

3.1 CLASSIFICATION APPROACH

The SVM is an administered classification system that limits an upper bound on its standard error. It endeavours to discover the hyperplane isolating two classes of information that will sum up best to future information.

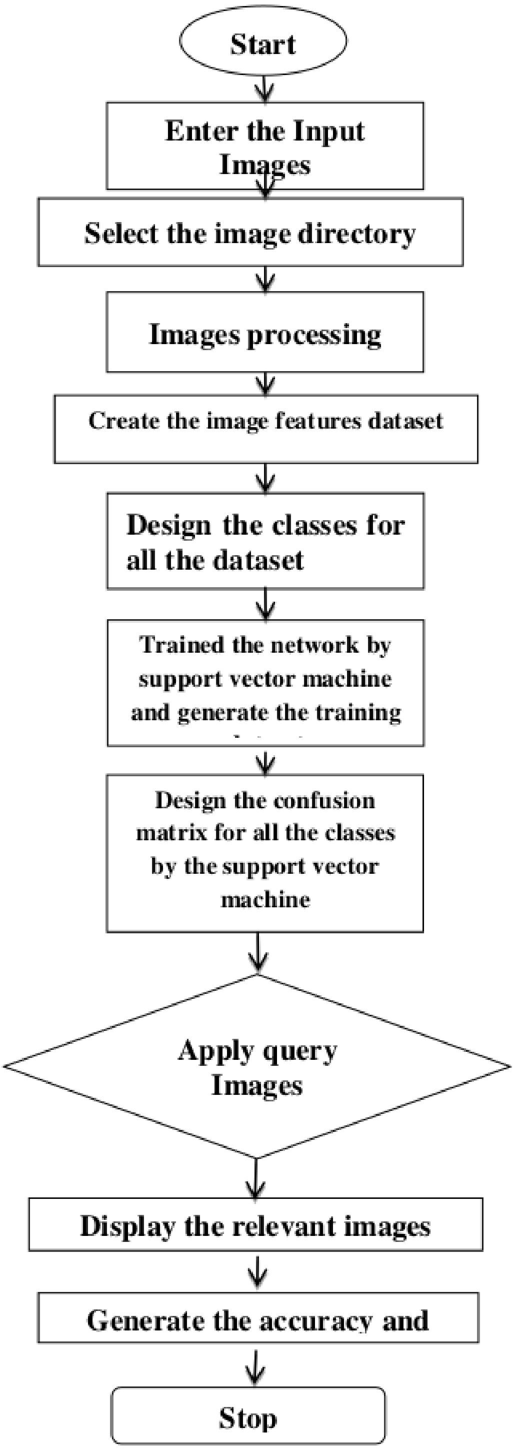

FIGURE 3. **The flow chart for the proposed model method.**

Such a hyperplane is the alleged greatest edge hyperplane that boosts the separation to the nearest focus from each class.

More concretely, mention the data points $\{X_0, \ldots \ldots X_N\}$ as well as class labels {y0...yN}, yi ∈ {−1, 1}, with any hyperplane extrication, the data classes which have the form in eqn (1)

$$y_i \left(w^T X_i + b\right) > 0 \ \forall_i \tag{1}$$

Let $\{w^T\}$ is used for all such hyperplanes. The concentrated margin hyperplane is well-defined in eqn (2)

$$w = \textstyle\sum_{i=0}^{N} \propto_i y_i X_i \ , \tag{2}$$

And the KKT conditions set b [6] where the $\{\alpha 0, \alpha 1, \ldots, \alpha N\}$ maximize-

$$L_D = \textstyle\sum_{i=0}^{N} \propto_i - \dfrac{1}{2} \textstyle\sum_{i=0}^{N} \sum_{j=0}^{N} \propto_i \propto_j y_i y_j X_i^T X_{j,} \tag{3}$$

Subject to.

$$\textstyle\sum_{i=0}^{N} \propto_i y_i = 0 \quad \propto_i \geq 0 \ \forall_i \quad . \tag{4}$$

For directly divisible information, just a subset of the $\alpha$is will be non-zero. These focuses are known as the help vectors, and all characterization performed by the SVM relies upon just these focuses and no others. In this manner, an indistinguishable SVM would result from a preparation set that precluded most of the rest of the models. This makes SVMs an alluring supplement to importance input: if the criticism framework can precisely recognize the basic examples that will end up being the help vectors, preparing time and naming exertion can, in the best case, be decreased definitely with no effect on the classifier exactness.

a) Performance Parameter

Evaluating the benefits and drawbacks of the image recovery system is crucial. The latest years have suggested A range of indices for measuring retrieval techniques. We use the most popular detail in this document and remember to assess the tests [31]. Precision relates to the proportion of the recovery amount of similar images TP to the entire amount of images TP + FP, which is used to assess the accuracy of the image recovery system Recall relates to the proportion of TP image resemblance scores in the register to the unlimited amount TP + FN used to evaluate the image recovery scheme. Eqn. 5 and 6 can quantify these two variables.

$$precision = \frac{TP}{TP + FP} \tag{5}$$

$$recall = \frac{TP}{TP + FN} \tag{6}$$

In other words, accuracy is the proportion of the retrieved images and all retrieved images. The recall is the proportion of the images extracted in the image database to all associated images. The more critical and more information from the images recovered is greater.

Accuracy and recall, nevertheless, are simultaneously limited. For instance, if the image database contains 1000 images, and only 100 images are relevant to the query image. If the first ten images are accurate, the precision is 1, whereas the recall value is only 0.1. Towards the other point, if all 1000 images were retrieved in the database, recall is 1, but precision is only 0.1. In this document, the precision-recall curve assessment is used to analyse the different recovery techniques, while calculating the extensive assessment index is used to measure the analysis output.

4. SIMULATION RESULTS AND ANALYSIS

In this chapter, we will explore the recovery efficiency for contrast with both the suggested template and the other designs. Furthermore, studies were carried out on three image database sets, demonstrating the suggested technique's legitimacy. The query image is an image input of the user, and the user wishes to use it as a sample to get pictures from the repository. The query image could have been from any source but not the repository. Fig. 4 shows an illustration of a query image.

The data set that trained the network had pre-processed pictures and worked with particular limitations. The image must, therefore, be pre-processed before the trained neural network analyzes it. The picture is transformed into a grey scale as a portion of the pre-processing phase and is resized to 28x28 pixels. Once the requested picture is a grayscale of 28x28 pixels, the qualified template can be used to evaluate it. Here provide a complete comparison with different datasets, the CIFAR-10 colour images [31], the WANG dataset [31], [32], and the Flower-Oxford [31] dataset.

FIGURE 4. **Query Image.**

FIGURE 5. **Sample Dataset of WANG.**

• **WANG dataset:** WANG dataset has 1000 images as the first test by the author [31]. There are ten types of these images, 100 of them each. The images in Fig. 5 are each category's depiction. Furthermore, these 1000 images can be split into two sections, 900 images will be used as teaching images, and the remaining 100 will be used as sample images. Of course, the ratio is equal for each image sort.

• **Dataset Oxford Flowers:** The second test uses the Oxford Flowers dataset [31]. The researchers selected an amount of very prevalent roses in the UK in this database, a total of 17 classifications. Eighty image representatives are selected for each group Fig. 6 shows some specimens. As with the first experiment, images will also be split into two classifications: teaching and testing, using the 60 images to practice and using the other 20 images to perform the test.

• **Dataset CIFAR-10:** As the final test, the data set CIFAR-10 [13], [14], [31] includes a 60,000 image serial of the 80 million small images. The 50000 and 100000 images were used for training and testing, respectively. Fig. 7 shows the sample image of the CIFAR dataset.

4.1 CNN RESULT ANALYSIS

A. CIFAR DATASET

Step 1: Select the datasets folder

Step 2: Extract the Feature and Dimension Reduction. Here image size is 64 x 64 bites.

Step 3: Apply the hash-deep CNN model. Used the Pre-trained VGG-16 model on a single CPU.

Step 4: Apply the query image for image retrieval.

Step 5: After applying the query image. Here carry out the retrieval by ranking its Hamming distances to the images from the folder. Here retrieve the top 16 images from the dataset.

B. FLOWER DATASET

Evaluating the benefits and drawbacks of the image recovery system is crucial. The latest years have suggested A range of indices for measuring retrieval techniques. We use the most popular detail in this document and remember to assess the tests. Precision relates to the proportion of the recovery amount of similar pictures TP to the unlimited amount of pictures TP + FP used to assess the accuracy of the image recovery system.

4.2 SUPPORT VECTOR MACHINE RESULT ANALYSIS

A. Wang Dataset

Step 1: Select the WANG dataset that has 1000 pictures. There are ten types of these pictures, 100 of them each. This dataset is categorized into ten different categories. Furthermore, these 1000 pictures can be split into two sections, 900 pictures will be used as teaching pictures, and the remaining 100 will be used as sample pictures. Of course, the ratio is equal for each picture sort.

Step 2: Select the query image for image retrieval and classification.

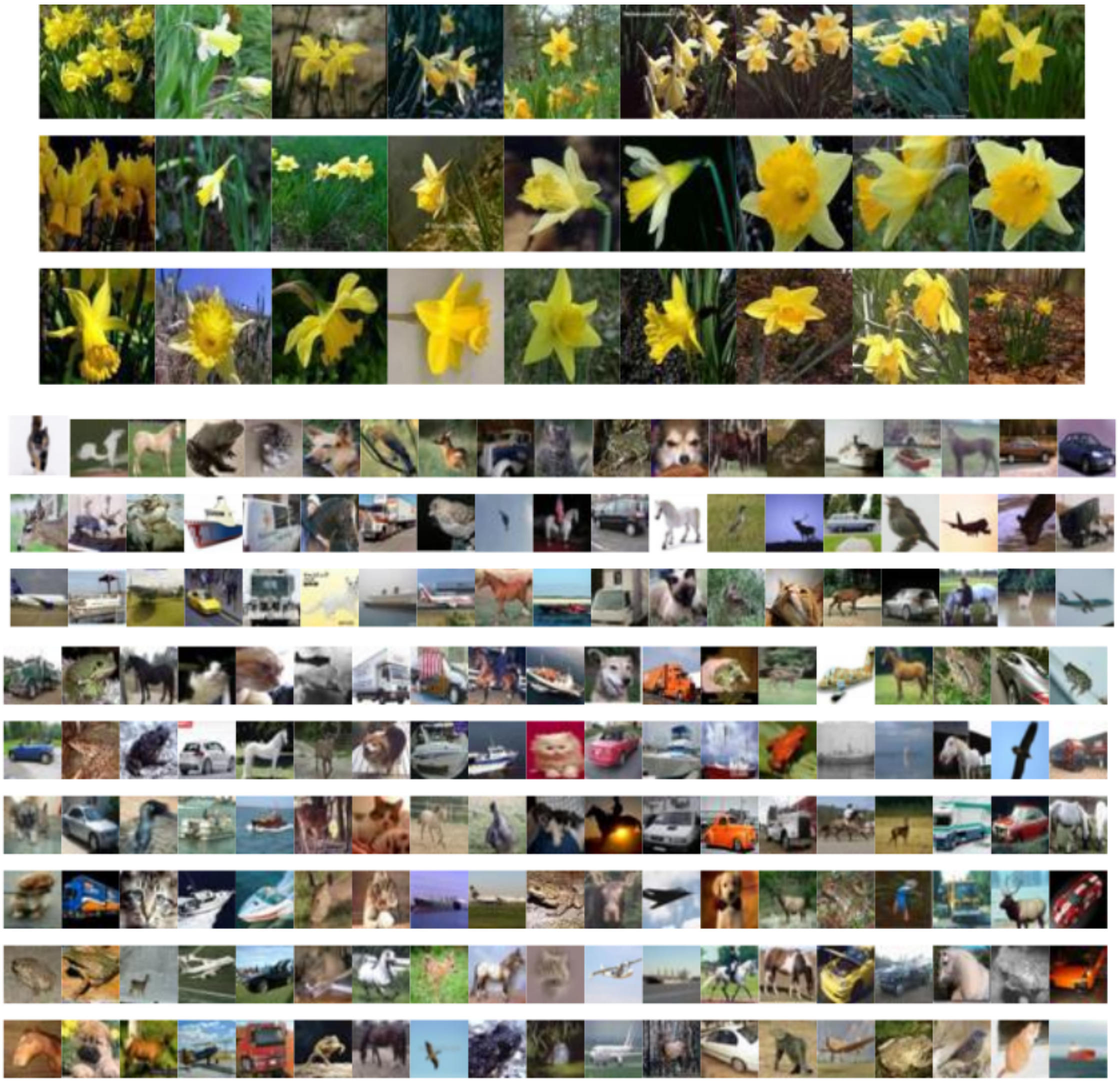

FIGURE 6. **Sample Dataset of Oxford Flower Dataset.**

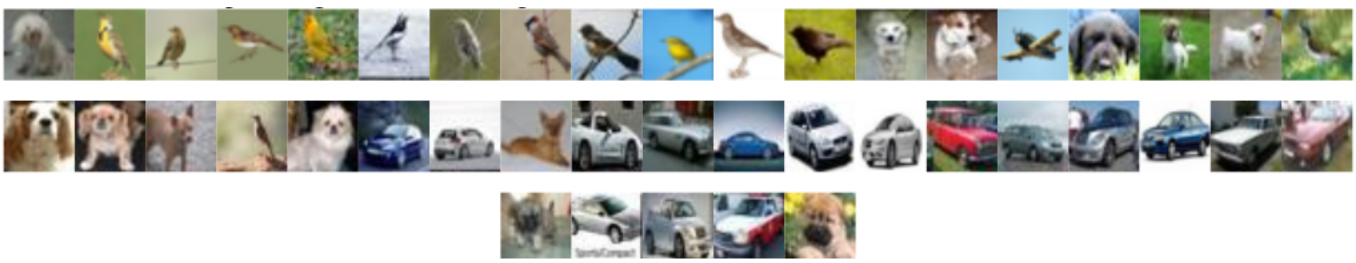

FIGURE 7. **Sample Dataset of CIFAR**

```
Training on single CPU.
Initializing image normalization.
|======================================================================================|
|     Epoch     |   Iteration   | Time Elapsed |  Mini-batch  |  Mini-batch  | Base Learning|
|               |               |  (seconds)   |     Loss     |   Accuracy   |     Rate     |
|======================================================================================|
|             1 |             1 |         1.95 |      -0.0000 |      100.00% |      1.00e-04 |
|            15 |            15 |        17.23 |      -0.0000 |      100.00% |      1.00e-04 |
|======================================================================================|

trainall =

     []

ENTER  QUERY IMAGE :'61.png'
```

FIGURE 8. **CNN screen output for CIFAR Dataset**

FIGURE 9. **query image for CIFAR Dataset.**

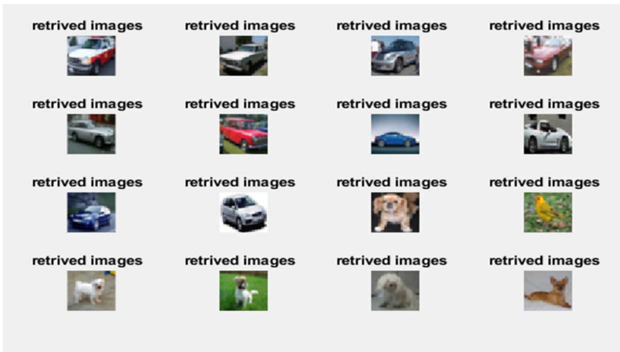

FIGURE 10. **Image retrieval for CIFAR Dataset.**

```
Training on single CPU.
Initializing image normalization.
|======================================================================================|
|     Epoch     |   Iteration   | Time Elapsed |  Mini-batch  |  Mini-batch  | Base Learning|
|               |               |  (seconds)   |     Loss     |   Accuracy   |     Rate     |
|======================================================================================|
|             1 |             1 |         1.99 |      -0.0000 |      100.00% |      1.00e-04 |
|            15 |            15 |        22.26 |      -0.0000 |      100.00% |      1.00e-04 |
|======================================================================================|

trainall =

     []

ENTER  QUERY IMAGE :'46.png'
```

FIGURE 11. **CNN screen output for Oxford Flower Dataset.**

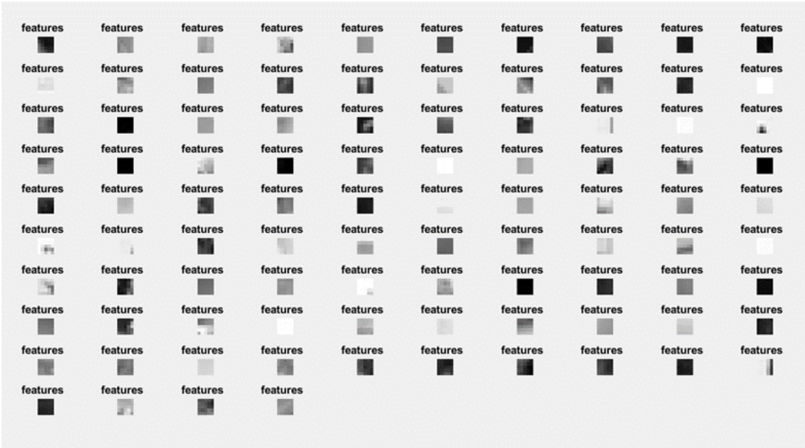

FIGURE 12. **Train network by CNN for Oxford Flower Dataset.**

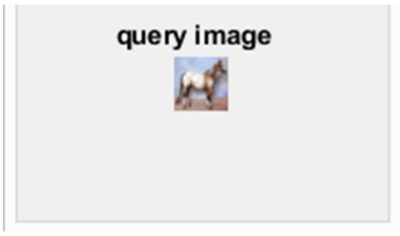

FIGURE 13. **query image for Oxford Flower Dataset**

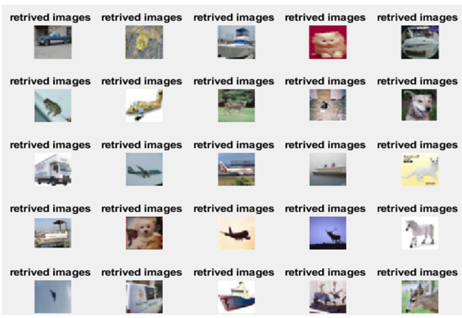

FIGURE 14. **CNN image retrieval for Oxford Flower Dataset.**

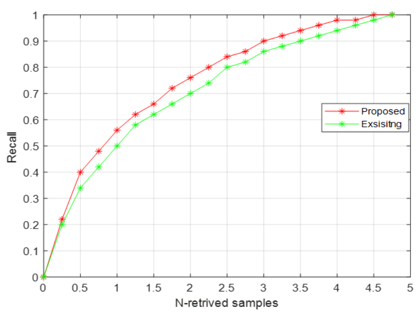

FIGURE 15. **Recall performance parameter comparison with the existing method.**

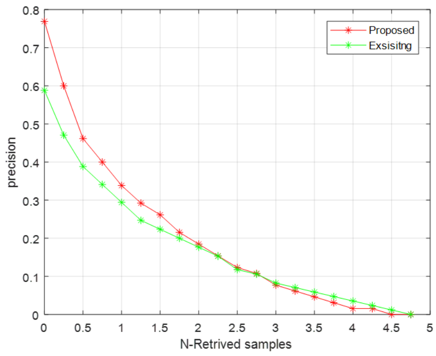

FIGURE 16. **Precision performance parameter comparison with the existing method**

Step 3: Design category set. So dataset belongs to this category for SVM classification. A category such as Africa, Beach, Monuments, Buses, Dinosaurs, Elephants, Flowers, Horses, Mountains, and Food.
Step 4: Extract query image features.
Step 6: Apply Gabor filters for the grayscale image.
Step 7: Construct the query Image feature vector.
Step 8: Choose your similarity metric and the number of returned images.
Step 9: Load the trained SVM dataset and extract image names in Figure 18.
Step 10: Construct labels.
Step 11: Divide the Dataset into split training/testing sets.
Step 12: Classify using the one-against-one approach, SVM with 3^{rd}-degree poly kernel.
Step 13: Used only training instances belonging to this pair.
Step 14: Design the confusion matrix in Figure 19.
Step 15: Measure the performance parameter such as precision, recall, and accuracy.
Step 16: Predicted Query Image Belongs to Class by SVM.

FIGURE 17. **query image for WANG Dataset.**

B. CIFAR DATASET

In other words, accuracy is the proportion of the retrieved images and all retrieved images. The recall is the proportion of the pictures extracted in the image database to all associated images. The more critical and more information from the images recovered is greater.

Accuracy and recall, nevertheless, are simultaneously limited. For instance, if the image database contains 1000 images, and only 100 images are relevant to the query image.

If the first ten images are accurate, the precision is 1, whereas the recall value is only 0.1. Towards the other point, if all 1000 images were retrieved in the database, recall is 1, but precision is only 0.1. In this document, the precision-recall curve assessment is used to analyze the different recovery techniques, while calculating the extensive assessment index is used to measure the analysis output. The proposed model shows the 70 -80 % improvement happening here.

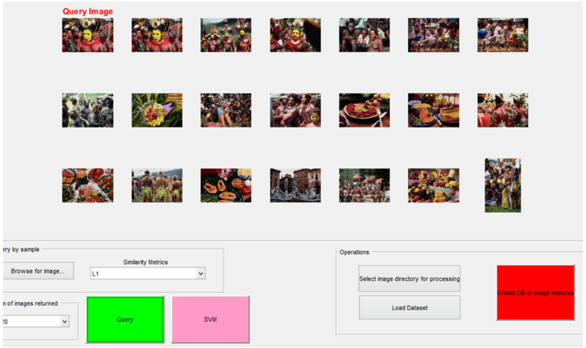

FIGURE 18. SVM image retrieval for WANG Dataset.

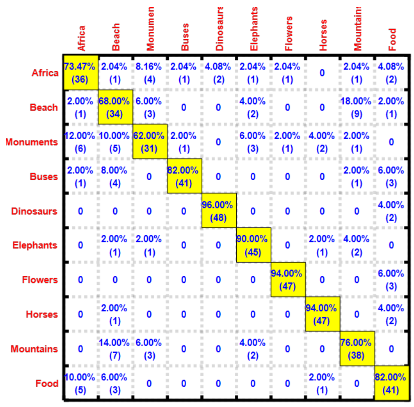

FIGURE 19. SVM confusion matrix for WANG Dataset.

FIGURE 20. query image for CIFAR Dataset.

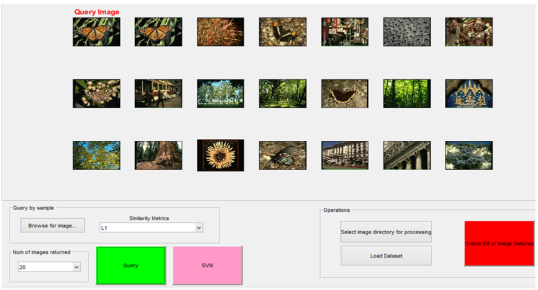

FIGURE 21. SVM image retrieval for flower Oxford Dataset

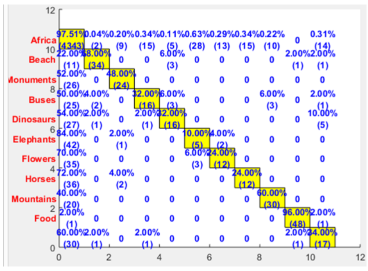

FIGURE 22. SVM confusion matrix for Oxford flower Dataset

Table 1. Accuracy comparison of the proposed model

Code Length	CIFAR-10	Oxford-Flower
64 bits	81.76 %	91.99 %

4.3 COMPARATIVE ANALYSIS

From Table .2, we observe that the existing method is not very sensitive to the length of the code. From 16 bits to 64 bits, the performances are just very marginal. The precision and recall of the proposed model are much higher than existing work.

Table 2. Precision Comparative analysis with existing work.

Code Length	CIFAR-10 (Existing)	CIFAR-10 Proposed Model
16 bits	0.3512	0.8181
32 bits	0.3791	0.8536
48 bits	0.3901	0.8958
64 bits	0.3840	0.87878

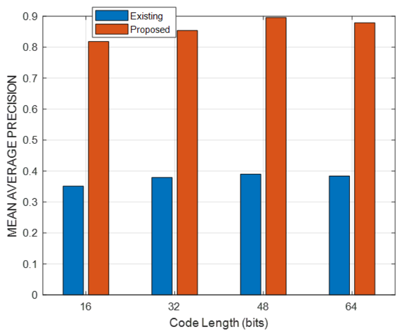

FIGURE 23. Precision Comparative analysis with existing work on CIFAR datasets.

5. CONCLUSION

This paper modified an unsupervised deep-hashing strategy by pseudo marks to address the current issue of scalable image retrieval with the help of a support vector machine classifier. Here incorporate maximum likelihood, variance and correlation, composed to discover virtual names & utilize the pseudo names as regulating data to prepare a deep system for the unsupervised hashing issue.

There, adaptive training through SVM can provide a guiding mechanism for looking for image repositories, surpassing the numerous traditional problems for refining plans. SVM not only achieves consistently high accuracy on a wide range of desired tests, but it does so easily. It retains high accuracy when required to express enormous amounts of images.

As seen plainly in the outcome page, the recall level and accuracy rate are increased after several training trials. This method can also be applied programmatically with an enticing aspect for image processing. It will be a great development in the world of computer science. And most significantly, reducing the overall time required may also be possible.

FUNDING

None

ACKNOWLEDGEMENT

None

CONFLICTS OF INTEREST

The author declares no conflict of interest.

REFERENCES

[1] Y. Liu, D. Zhang, G. Lu, and W. Y. Ma, “A survey of content-based image retrieval with high-level semantics,” *Pattern Recognit*, vol. 40, no. 1, pp. 262–282, 2007.

[2] Y. Liu, D. Zhang, and G. Lu, “Region-based image retrieval with high-level semantics using decision tree learning,” *Pattern Recognit*, vol. 41, no. 8, pp. 2554–2570, 2008.

[3] F. Bastien, “Theano: new features and speed improvements,” *ArXiv Prepr*, 2012.

[4] A. Brock, S. De, S. L. Smith, and K. Simonyan, “High-performance large-scale image recognition without normalization,” *International Conference on Machine Learning*, pp. 1059–1071, 2021.

[5] Y. Deng and B. S. Manjunath, “Unsupervised segmentation of color-texture regions in images and video,” *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 23, no. 8, 2001.

[6] Y. Han, F. Wu, Q. Tian, and Y. Zhuang, “Image annotation by input-output structural grouping sparsity,” *IEEE Trans. Image Process*, vol. 21, no. 6, pp. 3066–3079, 2012.

[7] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Commun. ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[8] L. Perez and J. Wang, “The effectiveness of data augmentation in image classification using deep learning,” *ArXiv Prepr*, 2017.

[9] C. F. Tsai, “Bag-of-words representation in image annotation: A review,” *Int. Sch. Res. Not*, vol. 2012, 2012.

[10] J. Wan, “Deep Learning for Content-Based Image Retrieval: A Comprehensive Study,” in *Proceedings of the 22nd ACM international conference on Multimedia*, pp. 157–166, ACM, 2014.

[11] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, and L. Fei-Fei, “ImageNet: A large-scale hierarchical image database,” in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255, IEEE, 2009.

[12] E. H. Hssayni, N. E. Joudar, and M. Ettaouil, “KRR-CNN: kernels redundancy reduction in convolutional neural networks,” *Neural Comput. Appl*, pp. 1–12, 2022.

[13] S. Divya, B. Adepu, and P. Kamakshi, “Image Enhancement and Classification of CIFAR-10 Using Convolutional Neural Networks,” *2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT)*, pp. 1–7, 2022.

[14] R. Doon, T. K. Rawat, and S. Gautam, “Cifar-10 classification using deep convolutional neural network,” *IEEE Punecon*, pp. 1–5, 2018.

[15] D. Agrawal, A. Agarwal, and D. K. Sharma, “Content-Based Image Retrieval (CBIR): A Review,” Recent Innov,” *Comput. Proc. ICRIC*, vol. 2, pp. 439–452, 2021.

[16] J. Choe, “Content-based image retrieval by using deep learning for interstitial lung disease diagnosis with chest CT,” *Radiology*, vol. 302, no. 1, pp. 187–197, 2022.

[17] R. Vishraj, S. Gupta, and S. Singh, “A comprehensive review of content-based image retrieval systems using deep learning and hand-crafted features in medical imaging: Research challenges and future directions,” *Comput. Electr. Eng*, vol. 104, pp. 108450–108450, 2022.

[18] M. Hemat, F. Shamsezat, M. Ezat, and K. Rafsanjani, “Image Retrieval based on Multi-features using Fuzzy Set,” *J. AI Data Min*, vol. 10, no. 4, pp. 569–578, 2022.

[19] X. Li, J. Yang, and J. Ma, “Recent developments of content-based image retrieval (CBIR),” *Neurocomputing*, vol. 452, pp. 675–689, 2021.

[20] T. Ji, X. Liu, C. Deng, L. Huang, and B. Lang, “Query-adaptive hash code ranking for fast nearest neighbor search,” *Proceedings of the 22nd ACM international conference on Multimedia*, pp. 1005–1008, 2014.

[21] X. Liu, L. Huang, C. Deng, B. Lang, and D. Tao, “Query-adaptive hash code ranking for large-scale multi-view visual search,” *IEEE Trans. Image Process*, vol. 25, no. 10, pp. 4514–4524, 2016.

[22] T. He, “Investigation of Animal Image Retrieval Algorithm Based on Deep Learning,” *Rev. Científica Fac. Cienc. Vet*, vol. 29, no. 1, pp. 128–138, 2019.

[23] E. S. S. H. Rizvi, *A New Framework to Bridge Semantic Gap for Semantics Based Image Retrieval*. 2014.

[24] N. Kumar, P. Rani, V. Kumar, S. V. Athawale, and D. Koundal, “THWSN: Enhanced Energy-Efficient Clustering Approach for Three-Tier Heterogeneous Wireless Sensor Networks,” *IEEE Sens. J*, vol. 22, no. 20, pp. 2022–2022.

[25] P. Rani and R. Sharma, “Intelligent transportation system for internet of vehicles based vehicular networks for smart cities,” *Comput. Electr. Eng*, vol. 105, pp. 108543–108543, 2023.

[26] G. Ansari, P. Rani, and V. Kumar, “A Novel Technique of Mixed Gas Identification Based on the Group Method of Data Handling (GMDH) on Time-Dependent MOX Gas Sensor Data,” in *Proceedings of International Conference on Recent Trends in Computing* (R. P. Mahapatra, S. K. Peddoju, S. Roy, , and P. Parwekar, eds.), vol. 600, pp. 641–654, Springer Nature, 2023.

[27] P. Rani, P. N. Singh, S. Verma, N. Ali, P. K. Shukla, and M. Alhassan, “An Implementation of Modified Blowfish Technique with Honey Bee Behavior Optimization for Load Balancing in Cloud System Environment,” *Wirel. Commun. Mob. Comput*, vol. 2022, pp. 1–14, 2022.

[28] P. Rani, S. Verma, S. P. Yadav, B. K. Rai, M. S. Naruka, and D. Kumar, “Simulation of the Lightweight Blockchain Technique Based on Privacy and Security for Healthcare Data for the Cloud System,” *Int. J. E-Health Med. Commun*, vol. 13, no. 4, pp. 1–15, 2022.

[29] L. Liu, M. Yu, and L. Shao, “Learning short binary codes for large-scale image retrieval,” *IEEE Trans. Image Process*, vol. 26, no. 3, pp. 1289–1299, 2017.

[30] X. Liu, J. He, and S. F. Chang, “Hash bit selection for nearest neighbor search,” *IEEE Trans. Image Process*, vol. 26, no. 11, pp. 5367–5380, 2017.

[31] H. Zhang, L. Liu, Y. Long, and L. Shao, “Unsupervised deep hashing with pseudo labels for scalable image retrieval,” *IEEE Trans. Image Process*, vol. 27, no. 4, pp. 1626–1638, 2017.

[32] R. Fu, B. Li, Y. Gao, and P. Wang, “Content-based image retrieval based on CNN and SVM,” *2016 2nd IEEE International conference on computer and communications (ICCC)*, pp. 638–642, 2016.