Research Journal of Computer Systems and Engineering

ISSN: 2230-8571, 2230-8563

Volume 02 Issue 01 - 2021 (January to June)

Page 16:22

Content Based Image Retrieval Based on Feature Extraction and Classification Using Deep Learning Techniques

Dr. Nitin Sherje
Associate Professor and Head of Department,
Smt. Kashibai Navale College of Engineering, Pune, India
npsherje@sinhgad.edu

Article History	***Abstract***
Received: 22 January 2021 Revised: 14 April 2021 Accepted: 19 May 2021	The Content based image retrieval plays a dynamic role in the contemporary scenario by being utilised to extract knowledge from photos. It is also a dynamic study area for many eras and is currently being rewarded more as a result of the theatrical and increase in the volume of digital photographs. The classifications of items based on their colour, texture, pattern, shape, layout, and location within the image, as well as other factors, are indexed and categorised according to the visual content of the image. Problem is identified in extraction of features and so the challenges are overcome by deep learning techniques. Initially the classification has been carried out using retrieval-based Inception V3-NET (RIV3-NET) algorithm. The noise has to reduce and enhance displacement with the smoothness by classifying invariant data of image using enhanced deep belief networks (EDBN). The simulation results show the enhanced retrieval of image and its parametric analysis. Keywords: Content based image retrieval, extraction of features, deep learning techniques, retrieval-based Inception V3-NET (RIV3-NET) algorithm, enhanced deep belief networks (EDBN).

1. Introduction

The image retrieval system performs the function of a classifier by categorising the images in the image database into two groups: relevant images and irrelevant images. Effective dimensionality reduction strategies would likewise be impossible with insufficient training data [1]. The retrieval procedure in a content-based image retrieval system uses content [2]. It involves looking for photos in a database and retrieving ones that appear visually similar to a given sample or test image [3]. The graphical components of an image, such as texture, colour, and spatial layout utilised to denote the image, are used in the process of content-based image retrieval, which additionally filters images based on their content to improve indexing and produce more accurate results[4].

The structure of this essay is as follows. We go over some related research, including both traditional CBIR technique formulations, in the next section. In Section 3 we discuss about the proposed methodology of the paper. Then in Section 4 we show results for CBIR. Then in Section 5 we conclude the paper with results obtained.

2. Literature Review

This problem is solved using a variety of methods, including the scale-invariant transform and the vector of locally aggregated descriptors. The computer part's next three approaches are then suggested

[5]. A hashing technique that extracts features from photos and discovers their binary representations was proposed in work [6]. The semantics assisted visual hashing method, proposed by author in [7], is an unsupervised visual hashing method (SAVH). Both offline learning and online learning are used in this method. In order to comply to the capabilities of CRB-CNN to extract the feature from only the visual input in the CBIR job, end-to-end tanning is applied at the very end without the use of any other metadata, annotations, or tags. The large-scale database image may also be retrieved using this method, and the retrieval performance was strong [8]. Hashing function receives effective attention in CBIR for efficient picture search [9]. The hashing function converts an image's high-dimensional visual data into a low-dimensional binary code that corresponds to the image's comparable content [10].

2.1 Research Methodology

This section discuss about the proposed methodology for the content based image retrieval.The above figure-1 shows proposed architecture of proposed methodology. Initially the input has been preprocessed and then feature is extracted in training phase. For testing phase the input is query image where it undergoes the preprocessing and feature extraction. Then the trained output and test output has been classified for identification of invariant data. There we obtain invariant data classification output. After this classified output we obtain retrieved image.

2.2 Feature extraction using retrieval-based inception V3-NET algorithm:

Increasing the depth and width of the network is a popular method for improving the performance of deep CNN models, however doing so will increase the number of network parameters. In order to circumvent the vanishing gradient problem, the Inception architecture also adds two extra softmax layers for forward propagation and employs an average pooling layer rather than a fully linked layer.

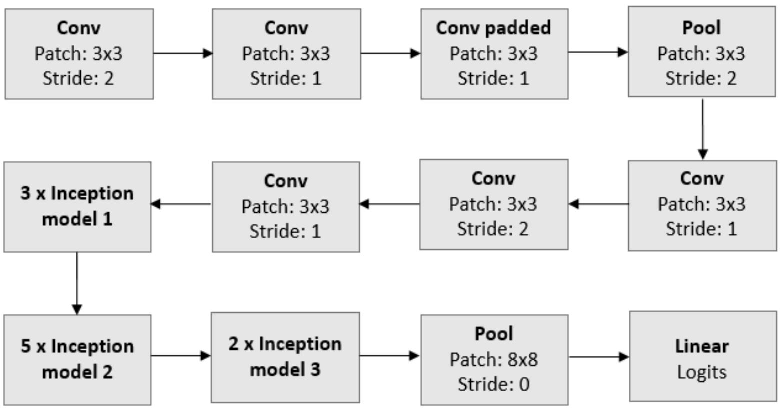

Figure 1: Inception module

As illustrated in Figure 1, Inception v3 simultaneously optimises the Inception network structure module using three distinct area grid sizes (35*35, 17*17, and 8*8).

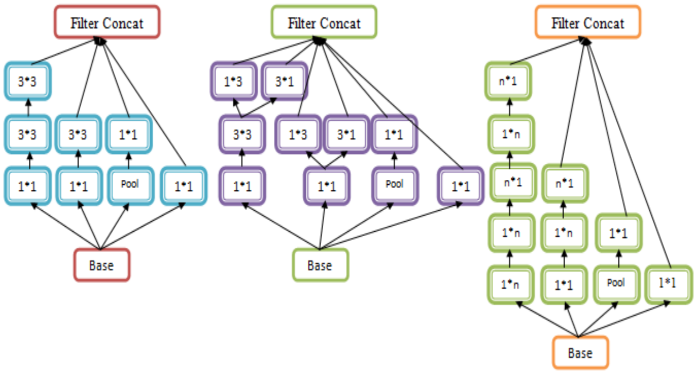

Figure 2: Inception module in RIV3-NET.

This procedure is repeated for the remaining feature patches, followed by applying this convolutional filter again and trying every possible match. The RIV3-NET network structure with CNN is depicted in figure 3.

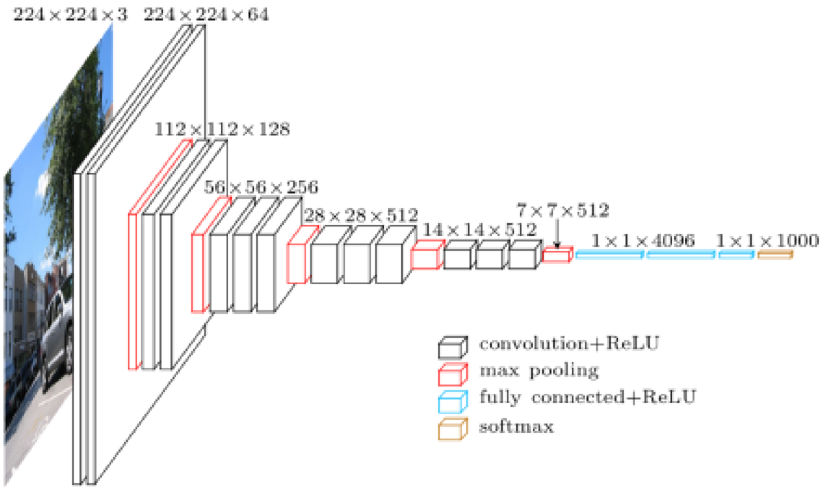

Figure 3: CNN Architecture

"Max pooling" refers to the CNN layer after that, which entails reducing the size of the picture stack. The Rectified Linear Unit (ReLU) method, commonly known as the normalisation layer of a CNN, entails setting all negative values in the filtered picture to 0. The ReLU layer then boosts the non-linear features of the model by repeating this step on all of the filtered images.

2.3 Invariant data classification using enhanced deep belief networks (EDBN):

Deep learning networks use several processing layers to describe high-level abstractions in data as a learning mechanism. First, the posterior probability distribution is used to compute the hidden layer values given the visible units eq. (1)-(3).

$$p(h_j = 1|v, \theta) = \sigma\left(a_j + \sum_{i=1}^{V} w_{ij} v_i\right) \quad [1]$$

$$p(v_j = 1|h, \theta) = \sigma\left(b_j + \sum_{j=1}^{H} w_{ij} h_j\right) \quad [2]$$

$$p(h_j = 1|v, \theta) = \mathcal{N}\left(b_j + \sum_{j=1}^{H} w_{ij}h_j , 1\right) \quad [3]$$

H is the quantity of concealed units. A Gaussian with mean and variance of and 2, $\mathcal{N}$ (μ, δ^2)Before using hidden unit likelihoods in (2) and (3), it's crucial to convert them to binary values (3). Third, the initial step is carried out once more using the values of the visible units that were rebuilt. The network weights will change after the procedure is finished by eq. (4),

$$\triangle\ w_{ij} \approx\ -\varepsilon(\langle v_i h_j\rangle_{data} - \langle v_i h_j\rangle_{recons}) \quad [4]$$

4. Performance Analysis

The performance analysis for the proposed RIV3-NET and EDBN is covered in this section. The picture retrieval from the input image using the existing and suggested methodologies is shown in Table 2. RIV3-NET performs better in image retrieval by producing output that is more efficiently optimised. The recall and precision of the system's retrieval performance can be evaluated. Precision examines the system's ability to recover only the relevant models, whereas recall evaluates the system's capacity to retrieve all relevant models.The metrics Precision, Recall, and F1 Score, which are specified below, are taken into account while assessing the performance of this suggested model.

Table-1 Analysis of Various Methods with Various Images

Sl.no	Input image	Methods	Precision (%)	Recall (%)	F1-score (%)	Accuracy (%)
1.	Ship	CNN	55	45	50	88
		KNN	56	48	52	89
		LBP	61	51	53	91
		RIV3-NET-EDBN	62	53	55	93
2.	Parachute	CNN	63	55	57	95
		KNN	65	56	61	91
		LBP	68	58	63	93
		RIV3-NET-EDBN	70	61	65	95
3.	Sign board	CNN	58	63	68	96
		KNN	62	65	71	97
		LBP	63	67	73	98
		RIV3-NET-EDBN	65	69	75	99

Table 1 displays the study of several techniques using different photos. It is demonstrated that the RIV3-NET outperforms the other approaches as the image size grows. The retrieved image and pertinent image are discussed. Ship, parachute, and sign board are among the images used as input.

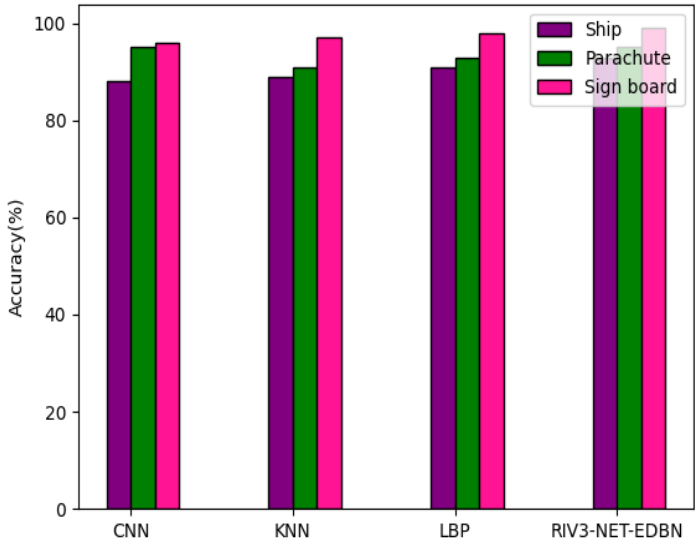

Figure 4: Comparison of Accuracy

The figure 4 shows comparison of accuracy between existing and proposed techniques. Graph has been plotted for the input image ship, parachute and signboard. The existing techniques compared are CNN, KNN, and LBP with proposed RIV3-NET with EDBN. When compared with existing techniques, the accuracy for RIV3-NET with EDBN has optimal accuracy in retrieving the image.

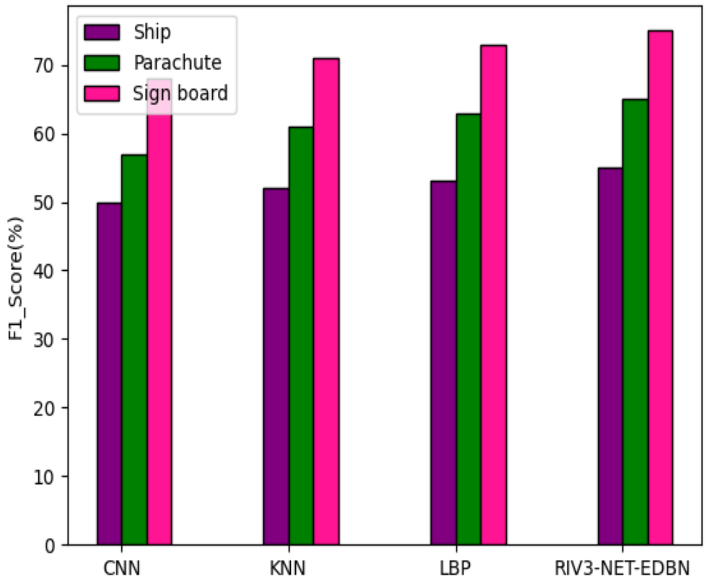

Figure 5: Comparison of F1- Score

Figure 5 shows comparison of accuracy between existing and proposed techniques. The existing techniques compared are CNN, KNN, and LBP with proposed RIV3-NET with EDBN. F1-Score for proposed techniques has been enhanced when compared with existing techniques.

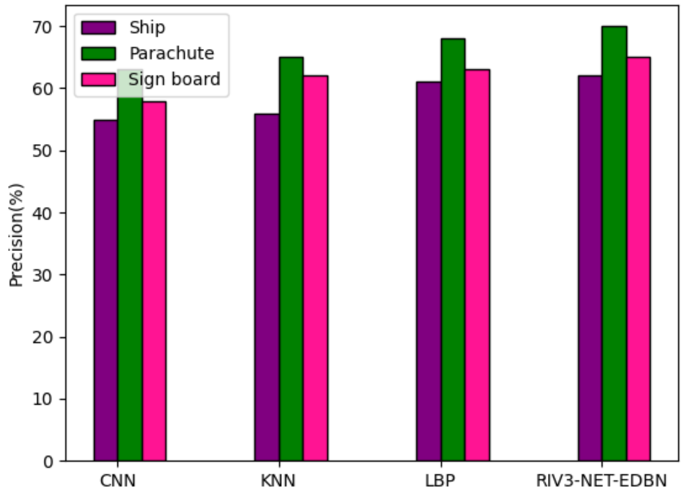

Figure 6: Comparison of Precision

The figure 6 shows comparison of accuracy between existing and proposed techniques. Graph has been plotted for the input image ship, parachute and signboard among the existing techniques CNN, KNN, LBP and proposed RIV3-NET with EDBN. Precision measures for proposed technique has been improved when compared with existing neural network techniques.

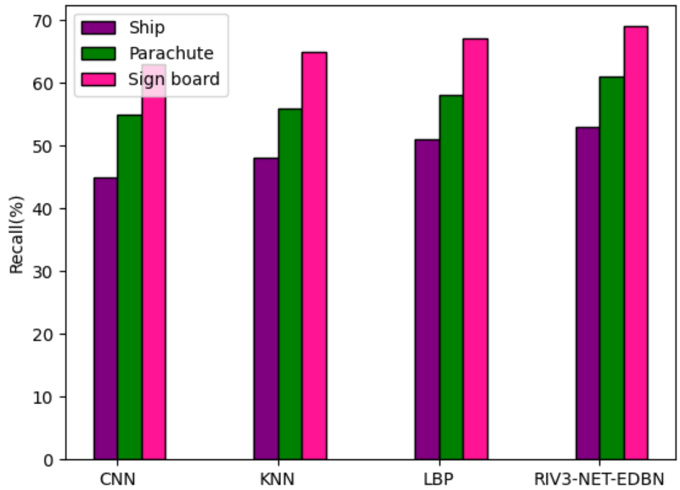

Figure 7: Comparison of Recall

The above figure 7 shows comparison of recall between existing and proposed techniques. The proposed RIV3-NET with EDBN shows the optimal results in comparison with existing CNN, KNN, and LBP.

5. Conclusion

This research proposed RIV3-NET with EDBN DL method for feature extraction. For efficient feature extraction in this instance, RIV3-NET with EDBN and chi square distance are used. The ship, parachute, and signboard images that are taken from the image frame yield the experimental results. This result is contrasted with those of the KNN (K-means Neural Network), Convolution Neural Network (CNN), and Local Binary Patterns (LBP) techniques that are currently in use. 90% accuracy was attained in recovering the image frames using the suggested technique.

Reference:

[1] Saritha, R. R., Paul, V., & Kumar, P. G. (2019). Content based image retrieval using deep learning process. *Cluster Computing*, *22*(2), 4187-4200.

[2] Mohamed, O., Khalid, E. A., Mohammed, O., &Brahim, A. (2017, October). Content-based image retrieval using convolutional neural networks. In *First International Conference on Real Time Intelligent Systems* (pp. 463-476). Springer, Cham.

[3] Alzu'bi, A., Amira, A., &Ramzan, N. (2017). Content-based image retrieval with compact deep convolutional features. *Neurocomputing*, *249*, 95-105.

[4] Rehman M, Iqbal M, Sharif M and Raza M, "Content based image retrieval: survey", *World Applied Sciences Journal*, vol.19, no.3, pp.404-412.

[5] Kondylidis N, Tzelepi M and Tefas A, "Exploiting tf-idf in deep convolutional neural networks for content based image retrieval", *Multimedia Tools and Applications*, vol.77, no.23, pp.30729-30748, 2018.

[6] Shi X, Sapkota M, Xing F, Liu F,Cui L and Yang L, "Pairwise based deep ranking hashing for histopathology image classification and retrieval", *Pattern Recognition*, vol.81, pp.14-22, 2018.

[7] Zhu L, Shen J, Xie L and Cheng Z, "Unsupervised visual hashing with semantic assistant for content-based image retrieval", *IEEE Transactions on Knowledge and Data Engineering*, vol.29, no.2, pp.472-486, 2017.

[8] Alrahhal, M., &Supreethi, K. P. (2019, February). Content-based image retrieval using local patterns and supervised machine learning techniques. In *2019 Amity International Conference on Artificial Intelligence (AICAI)* (pp. 118-124). IEEE.

[9] Xu, H., Wang, J. Y., & Mao, L. (2017, June). Relevance feedback for Content-based Image Retrieval using deep learning. In *2017 2nd International Conference on Image, Vision and Computing (ICIVC)* (pp. 629-633). IEEE.

[10] Liu, P., Guo, J. M., Wu, C. Y., &Cai, D. (2017). Fusion of deep learning and compressed domain features for content-based image retrieval. *IEEE Transactions on Image Processing*, *26*(12), 5706-5717.