International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Novel Technique in Content Based Image Retrieval using Classification by Deep Learning in Artificial Intelligence

Nilanjana Saha [1], P Thiruvannamalai sivasankar [2], Ms.Rahama Salman [3], Badria Sulaiman Alfurhood [4], Mr. Aishwary Awasthi [5], Sukhwant Singh Bindra [6]

Submitted: 18/08/2022 **Accepted:** 23/11/2022

Abstract:A popular technique for retrieving images from sizable and unlabeled image collections is content-based image retrieval (CBIR). One of the subcategories of the soft computing phenomenon known as "deep learning" allows for the retrieval of data from millions of separated images. This research propose novel technique in comprehensive description for content based image retrieval by classification utilizing DL techniques. Here input image has been collected and processed for noise removal. The processed image has been extracted and classified using FFCNN. Based on extracted features by calculating the similarity index of the images based on ranking matrix and distance calculation the image has been retrieved. The experimental analysis has been carried out in terms of accuracy, precision, recall, RMSE. the proposed technique attained accuracy of 95%, precision of 79%, recall of 68% and RMSE of 61%.

Keywords: *Content-based image retrieval, classification, DL, noise removal, ranking matrix, distance calculation*

1. Introduction

Along with Bing photo search, search engines like the CBIR engine from Microsoft (Public Company), the CBIR machine from Google, CBIR search engine, Gazopa, Imense Image Search Portal, and others appear to have grown quickly in recent years. The task of retrieving photos has also proven to be difficult for Com [1]. In essence, the database images are ranked in decreasing order of similarity based on how much they resemble the query image. The method used to calculate the similarity score between two photos should ideally be discriminative, reliable, and effective. Deep learning is one of the subcategories of the soft computing phenomenon, and it may be used to extract data from millions of separated images [2]. The feature representation and similarity measurement, which have been thoroughly explored by multimedia researchers for decades, are key components that determine how well a content-based image retrieval system performs while retrieving images. In essence, a feature descriptor represents the content of the photos (i.e., the visual appearance) in terms of colour, texture, shape, gradient, etc. Similarity between the photos is determined by comparing the feature vectors of the respective photographs. Thus, the feature descriptor representation of the image has a significant impact on how well any CBIR method performs [3].

The contribution of this research is as follows:

To propose novel technique in comprehensive description for content based image retrieval by classification utilizing DL techniques.

The input image has been collected and processed for noise removal. The processed image has been extracted and classified using FFCNN.

2. Literature Works

There were numerous local and global elements to express the image qualities and content for various earlier content-based techniques [4]. For the extraction of colour and texture features in [5], primitives and colony filters are utilised. In their work, an image is broken up into numerous smaller pieces, and the colour moments of each block are retrieved according to an existing method [6]. Using a clustering method and a predetermined colour feature vector approach that are computed from the query image and the photos in the image database, these moments are grouped into several classes [7]. Each digital value will serve as a representation of the distance between each digital image, however up until recently, we were unable to extract an exact value from each of the images we searched specifically [8]. The method of communication used between the point nodes of the image is defined in some studies. The precision is 59.61 on average. By extracting the object features from the photos, object-based image retrieval systems get images from a database. This technique compares each segmented zone in the database picture to a region in the user-provided query image [9]. These

[1] *Amity Institute of International Studies, Research Scholar Amity University, Noida.*
sahanilanjana.05@gmail.com
[2] *Professor, Department of Computer Science and Engineering, Jain(Deemed-to-be University), Bangalore, India.*
tp.sivasankar@jainuniversity.ac.in
[3] *Lecturer, Department of Information Technology and Security, College Of CSIT, Jazan University, Jazan,KSA.*
rabdol@jazanu.edu.sa
[4] *Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Saudi Arabia.*
bsalfurhood@pnu.edu.sa
[5] *Research Scholar, Department of Mechanical Engineering, Sanskriti University, Mathura, Uttar Pradesh, India.*
aishwary@sanskriti.edu.in
[6] *Amity Institute of International Studies, Research Director (AIIS & AIPP), Amity University, Noida.*
ssbindra@amity.edu

image retrieval techniques typically work well for items that stand out against the background and have unique colours or textures. A model that integrates colour histograms and colour moment feature extraction techniques was proposed. They claimed that index sorting outperformed other methods. A model based on colour and texture attributes is proposed by [10].

3. Comprehensive Description on Artificial Intelligence

The potentialization of numerous aspects of daily life, including customer service, finance, sales and marketing, administration, and technical operations in various industries, is facilitated by the application of artificial intelligence (AI) in business. It is crucial to keep in mind that it has not developed to take over human jobs, but to complement them and enable individuals to fully express their potential and creativity.

Proposed Content-based image retrieval:

This section discuss novel technique in comprehensive description for content based image retrieval by classification utilizing DL techniques. here the input image has been collected and processed for noise removal. The processed image has been extracted and classified using feed forward convolutional neural network. The overall proposed architecture is shown in figure-1.

A contrast function is defined in terms of a fairly straightforward measurement of the dependency of a collection of random variables. For the sake of simplicity, consider the situation of two univariate random variables x1 and x2, and assume that F is a vector space of functions from R to R by eq.(1).

$$\rho_{\mathcal{F}} = \max_{f_1,f_2\in\mathcal{F}} \operatorname{corr}\left(f_1(x_1), f_2(x_2)\right) = \max_{f_1,f_2\in\mathcal{F}} \frac{\operatorname{cov}\left(f_1(x_1), f_2(x_2)\right)}{\left(\operatorname{var} f_1(x_1)\right)^{1/2}\left(\operatorname{var} f_2(x_2)\right)^{1/2}} \quad (1)$$

Clearly, F-correlation is equal to 0 if variables x1 and x2 are independent. Opposite is also true if set F is sizable enough. For instance, it is commonly known that the statement ρF = 0 denotes the independence of x1 and x2 if F comprises the Fourier basis. Assuming that F is an RKHS on R, K(x, y) is associated kernel, and (x) = K(, x) is feature map eq.(2).

$$f(x) = \langle \Phi(x), f\rangle, \ \forall f \in \mathcal{F}, \forall x \in \mathbb{R}.$$

$$\operatorname{corr}\left(f_1(x_1), f_2(x_2)\right) = \operatorname{corr}\left(\langle \Phi(x_1), f_1\rangle, \langle \Phi(x_2), f_2\rangle\right)$$

$$D(H \circ \mu_X)(0)\bar{\Omega} = \frac{d}{d\varepsilon}(H \circ \mu_X)(\varepsilon\bar{\Omega})\Big|_{\varepsilon=0}$$

$$= \sum_{1\le i<j\le m}^{m} \omega_{ij}\left(\tau_{ij} - \tau_{ji}\right),$$

$$\tau_{ij} = \sum_{r=1;r\neq i}^{m} \mathbb{E}_{k,l}\left[\phi'\left(x_i^T \bar{w}_{kl}\right) x_j^T \bar{w}_{kl} \phi\left(x_r^T \bar{w}_{kl}\right)\right] + \mathbb{E}_{k,l}\left[\phi'\left(x_i^T \bar{w}_{kl}\right) x_j^T \bar{w}_{kl}\right] \mathbb{E}_{k,l}\left[\phi\left(x_r^T \bar{w}_{kl}\right)\right] - 2\mathbb{E}_k\left[\mathbb{E}_l\left[\phi'\left(x_i^T \bar{w}_{kl}\right) x_j^T \bar{w}_{kl}\right] \mathbb{E}_l\left[\phi\left(x_r^T \bar{w}_{kl}\right)\right]\right] \quad (2)$$

Feed Forward Convolutional Neural Network Based Classification:

A parallel-operating feedforward neural network features a multilayer architecture with interconnected processing units. Connections between parts play a significant role in how well the network functions. The network's adaptive parameters for the given input vector p are made up of the weights (w) and biases (b) combined. The network can learn both linear and nonlinear relationships between input and output vectors thanks to its multiple layers of nonlinear transfer function neurons. Network must be trained to strike a compromise between being able to react appropriately to input patterns and giving good responses. The i-th neuron of k-th layer's input signal is given by eq.(3):

$$s_i^k = \sum_{j=1}^{n} w_{ij}^k y_j^{k-1} \quad (3)$$

where y k1 j is output signal value of layer j's previous neuron and w k ij is the weight of link between neuron I of layer k and neuron j of layer preceding. The ith neuron's output signal is given by eq.(4):

$$y_i^k = f\left(s_i^k\right) = f\left(\sum_{j=1}^{n} w_{ij}^k y_j^{k-1}\right) \quad (4)$$

Bipolar linear function is one of the most widely utilised. The formula for it is given by eq. (5)

$$f\left(s_i^k\right) = \frac{2}{1+e^{-\alpha s_i^k}} - 1 = \frac{1-e^{-\alpha s_i^k}}{1+e^{-\alpha s_i^k}} \quad (5)$$

Equation $E = \sum_T \sum_{i=1}^{n} (d_i - y_i)^2$. can be used to characterise the total error of the training set using MSE. The weight is adjusted using the following eq. (6):

$$w_{ij}^k = w_{ij}^k + \eta \nabla w_{ij}^k$$

$$\nabla w_{ij}^k = \frac{\partial E}{\partial w_{ij}^k} = \frac{1}{2} \cdot \frac{\partial E}{\partial s_i^k} \cdot 2 \cdot \frac{\partial s_i^k}{\partial w_{ij}^k} = 2\delta_i^k y_j^{k-1},$$

$$\delta_i^K = \frac{1}{2} \cdot \frac{\partial E}{\partial s_i^k} = \frac{1}{2} \cdot \frac{\partial \left(d_i^K - y_i^K\right)^2}{\partial s_i^k} = f'\left(s_i^K\right) \cdot \left(d_i^K - y_i^K\right)$$

$$\delta_i^k = \frac{1}{2} \cdot \frac{\partial E}{\partial s_i^k} = \frac{1}{2} \cdot \sum_{j=1}^{N_{k+1}} \frac{\partial E}{\partial s_j^{k+1}} \frac{\partial s_j^{k+1}}{\partial s_i^k} = f'\left(s_i^k\right) \sum_{j=1}^{N_{k+1}} \delta_j^{k+1} w_{ij}^{k+1} \quad (6)$$

Nk+1 is the number of neurons on the (k + 1) th layer. It makes use of an array called D that has values. This crooked array has L rows and as many columns in a row as the layer's total number of neurons. The alternate syntax for writing [i][j] is Ai, which is used for all array items. Additionally, the for-loops' employed numerical intervals are half-open. Additionally, there are symbols like LR, s, y ki, and w ki j by eq. (7).

$$d_{st} = 1 - \frac{(x_s - x'_s)(x_t - x'_t)'}{\sqrt{(x_s - x'_s)(x_s - x'_s)'}\sqrt{(x_t - x'_t)(x_t - x'_t)'}} \quad (7)$$

$$x'_i > x'_j \Leftrightarrow f(x'_i) f\left(x'_j\right)[10]$$

While indexing correlation of pixel co-ordinates is used to determine ranking. (x'_i, x'_j) has been frequently utilised for classification on pairs of pixel values (x'_i, x'_j) rather than using either the best or poorly ranked. It is treated as learning issue and it is interpreted from ranking issue. The new vector (x'_i, x'_j) provided by the more established relation (x'_i, x'_j) such that by eq.(8)

$$\left(x'_i - x'_j, z = \begin{Bmatrix} +1 & y_i > y_j \\ -1 & y_j > y_i \end{Bmatrix}\right) \quad (8)$$

4. Experimental Analysis

Each network was trained over the course of 100 epochs, with each epoch consisting of 50 iterations and a batch size of 1. We employed a planned training where the learning rate was set for the first 60 epochs at 0.005, the following 20 till 80 at 0.001, and the final 20 at 0.0001. A computer with an Intel Core i7-8700 processor and 15.6 GiB of RAM was used to train the models. Any particle shape, regardless of size or aspect ratio, can be trained using the given settings. A high resolution digital camera was used to capture one-view photos of apples, oranges, potatoes, and tomatoes in a diffusely lit setting. Variety of fruits and vegetables were chosen over those that were only one colour since their bi-colored skin poses additional segmentation

challenges because of colour transition zones.

Table-1 Comparative analysis between proposed and existing technique

Parameters	OIRSRI	CTF	CD_AI_CBIR_DL
Accuracy	89	92	95
Precision	71	73	79
Recall	61	63	68
RMSE	55	59	61

The above table-1 shows comparative analysis between proposed and existing technique in terms of accuracy, precision, recall and RMSE. Here the existing technique compared are OIRSRI and CTF with proposed technique.

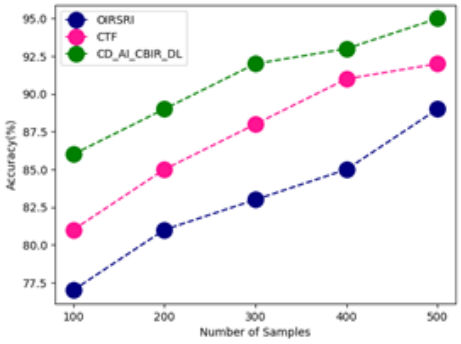

Figure-2 Comparison of Accuracy

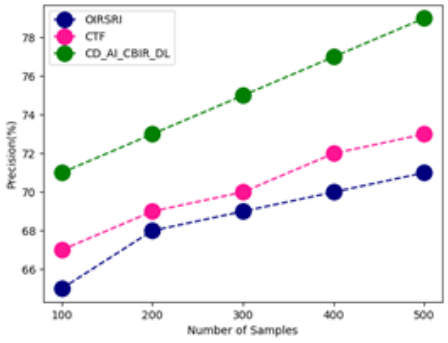

Figure-3 Comparison of Precision

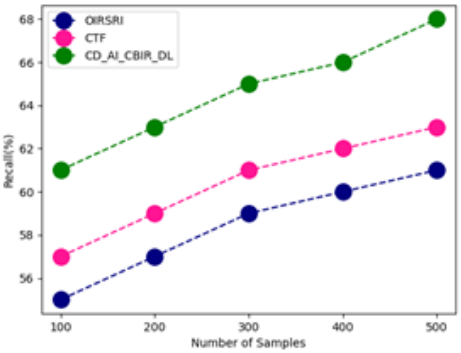

Figure-4 Comparison of recall

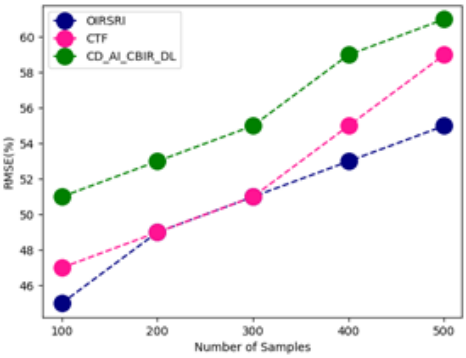

Figure-5 Comparison of RMSE

From above figure 2-5 the comparative analysis between proposed and existing technique has been shown in terms of accuracy, precision, recall and F_1 score. Here the proposed technique attained accuracy of 95%, precision of 79%, recall of 68% and RMSE of 61%. One specification for assessing classification methods is accuracy. Official description of accuracy is as follows: Total number of accurate guesses is equal to total number of accurate guesses. By dividing number of accurate predictions by overall sample size, we may determine accuracy. Our model was 44 percent accurate on this multiclass problem, according to the outcome. One indicator of the model' s performance is precision, or the quality of a successful prediction. Total number of accurate positive predictions is divided by total number of real positives to determine precision. Recall literally refers to how many right hits were also discovered, or how many genuine positives were remembered. Precision is the percentage of returning hits that were true positive, or correct hits. The recall is determined by comparing the proportion of correctly labelled Positive samples to all Positive samples. Model ability to make distinctions Recall is used to measure positive samples. As more positive samples are identified, the recall rises.

5. Conclusion

This research propose novel technique in comprehensive description for content based image retrieval by classification utilizing DL techniques. The processed image has been extracted and classified using FFCNN. The adoption of technologies like AI at all levels and operations of businesses will become a reality to boost their competitiveness over the coming years, and digital endeavours won't be isolated projects or initiatives in firms anymore. AI is starting to permeate company operations. Thus, the similarity computation between images determines how well any image retrieval algorithm performs. the proposed technique attained accuracy of 95%, precision of 79%, recall of 68% and RMSE of 61%.

Reference

[1] Bartłomiej Mellera, Kamil Matula, "Image classification with Feed-Forward Neural Networks", Symposium for Young Scientists in Technology, Engineering and Mathematics, Online, May 20 2020.

[2] Garg, M., & Dhiman, G. (2021). A novel content-based image retrieval approach for classification using GLCM features and texture fused LBP variants. Neural Computing and Applications, 33(4), 1311-1328.

[3] Saleem, S., Amin, J., Sharif, M., Anjum, M. A., Iqbal, M., & Wang, S. H. (2022). A deep network designed for segmentation and classification of leukemia using fusion of the transfer learning models. Complex & Intelligent Systems, 8(4), 3105-3120.

[4] Jaiswal, A., Gianchandani, N., Singh, D., Kumar, V., & Kaur, M. (2021). Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning. Journal of Biomolecular Structure and Dynamics, 39(15), 5682-5689.

[5] Yu, H., Yang, L. T., Zhang, Q., Armstrong, D., & Deen, M. J. (2021). Convolutional neural networks for medical image analysis: state-of-the-art, comparisons, improvement and perspectives. Neurocomputing, 444, 92-110.

[6] Shankar, K., & Perumal, E. (2021). A novel hand-crafted with deep learning features based fusion model for COVID-19 diagnosis and classification using chest X-ray images. Complex & Intelligent Systems, 7(3), 1277-1293.

[7] Bansal, M., Kumar, M., & Kumar, M. (2021). 2D object recognition: a comparative analysis of SIFT, SURF and ORB feature descriptors. Multimedia Tools and Applications, 80(12), 18839-18857.

[8] El Asnaoui, K., Chawki, Y., & Idri, A. (2021). Automated methods for detection and classification pneumonia based on x-ray images using deep learning. In Artificial intelligence and blockchain for future cybersecurity applications (pp. 257-284). Springer, Cham.

[9] Wang, Z., Zhao, W., Du, W., Li, N., & Wang, J. (2021). Data-driven fault diagnosis method based on the conversion of erosion operation signals into images and convolutional neural network. Process Safety and Environmental Protection, 149, 591-601.

[10] Ibrahim, A. U., Ozsoz, M., Serte, S., Al-Turjman, F., & Yakoi, P. S. (2021). Pneumonia classification using deep learning from chest X-ray images during COVID-19. Cognitive Computation, 1-13.