Original Article

DEEP LEARNING TECHNIQUES FOR EXOTICISM MINING FROM VISUAL CONTENT BASED IMAGE RETRIEVAL

Dr. S.M.D.Mathuravalli[1] Narayanansamy Rajendran[2], Dr.K.Bagyalakshmi[3], Dr. Dilip R[4], Dr Anuradha Ranjan[5], Dr. Ipsita Das[6], Dr. Amit Chauhan[7]

[1]Associate Professor, Department of Food Science and Nutrition, Holy Cross Home Science College, Thoothukudi-628 003. Tamil Nadu.
E-mail: smdmathuravalli@gmail.com
[2] Lecturer, Information Technology Department, University of Technology and Applied Sciences – Nizwa.
E-mail: naren.rajendran@nct.edu.om
[3] Assistant Professor, Department of Electrical and Electronics Engineering, Sri Ranganathar Institute of Engineering and Technology.
E-mail: bagyalakshmi@sriet.ac.in
[4]Assistant Professor, Department of Electronics and Communication Engineering, Dayananda Sagar Academy of Technology & Management, Udayapura, Bengaluru-560082, India. E-mail: dilipr-ece@dsatm.edu.in
[5]Assistant professor , Department of Psychology, Amity University, Mohali, Punjab-140306, India.
E-mail: anuradha.ranjan24@gmail.com
[6]Assistant Professor II, KIIT School of Law, KIIT Deemed to be University, Prasanti Vihar, Patia, Bhubaneswar, Odisha.
E-mail: Ipsitadas50@gmail.com, Ipsitadas@kls.ac.in
[7]Assistant professor, Department of Life sciences, CHRIST (Deemed to be university), Bangalore, Karnataka-560029, India.
E-mail: amit_chauhan777@yahoo.in

DOI: 10.47750/pnr.2023.14.S01.130

Abstract

Early diagnosis has been aided by electronic restorative imaging and examination approaches employing several modalities. The advancement of computer-aided retrieval systems in recent years has made them a nondestructive and popular tool for disease identification in medical photographs. In this paper, an adaptable Gabor wavelet filter bank and a feature descriptor based on Texton are created for medical picture retrieval. The suggested descriptor basis's architecture allows for flexibility in extracting the dominating directional characteristics from medical pictures. In addition, we provide an unique end-to-end integrated deep learning model that employs the Convolutional Neural Network (CNN) and the Long Short-Term Memory cell (LSTM). Using datasets such as New Caltech, Corel-1000, Oliva, and Corel-10,000, the proposed integrate deep learning descriptor is compared to existing descriptors such as CCM, CHD, MTH, and MSD.

Keywords: Image processing , Convolutional Neural Network , Texture, Feature Selection, Pattern Recognition.

1. Introduction

Continuous advancements in PC technology have enabled the development of many types of totally automated computer-based diagnostics. Presently, restorative picture investigation is an exploration zone that pulls in a tonne of anxiety from the two researchers and physicians. Early detection has been encouraged by electronic restorative imaging and examination approaches employing several modalities [1][2]. Using Scheie's arrangement scheme, the presence of haemorrhages is widely employed to examine DR or hypertensive retinopathy. Notwithstanding recognising microaneurysms, it is problematic for ophthalmologists to locate them in noncontrast fundus photographs. Because the unpredictability of a microaneurysm image is quite low, ophthalmologists often identify

microaneurysms using fluorescein angiograms. In any event, it is difficult to use fluorescein as a different technique of diagnosis for all of the medical examinees subjected to mass screening.

Feature extraction and similarity assessment are standard components of a medical image retrieval system. This research focuses on the feature extraction stage which is the cornerstone for a successful picture retrieval system. Several approaches have been presented for the extraction of characteristics from the medical pictures. Local binary pattern (LBP) [3] has a low computing complexity and works well in texture classification of medical photos. Furthermore, several LBP variants are proposed, including centre symmetric LBP [4, data driven LBP [5, local ternary patten (LTP) [6, and centre symmetric LTP,MTH, and MSD [7-9], in which a relationship is established between the centre pixel and its neighbouring pixels for medical image retrieval images.

Contours often include important visual information about a picture. Contours have been extensively employed in various practical applications in computer vision. Despite the fact that several contour detection algorithms have been developed over the last few decades, contour detection remains a difficult issue in the picture field. Among the non-learning techniques, several early systems, such as the renowned Canny detector, locate contours by removing edges where the brightness or colour changes abruptly. However, such algorithms often use normal kernels, such as the Gaussian and Gabor filters, to assess the extents of local changes, and hence cannot cope with textures. To overcome this issue, numerous texture suppression approaches have been developed. Methods based on nonclassical receptive field inhibition, sparseness measurements, surround-modulation, and other techniques are examples. Texture suppression has been shown to increase contour detection performance. Nonetheless, these approaches still largely exploit low-level local characteristics. Moreover, some of them are computationally demanding, which leads to challenges in actual applications[10].

Because of its superior soft-tissue contrast and accessibility of multispectral pictures, medical imaging is a prominent non-invasive tool for the visualisation of various disorders in the brain. CAD systems have been created to aid clinicians in quick diagnosis by using information from MR images. Based on the precise qualities contained in the medical photos, CAD systems may deliver a diagnosis. For classifying normal/abnormal brain MR images, these systems typically utilise the phases of preprocessing, attribute extraction, selection, and classification. Numerous strategies for detecting aberrant brain pictures using traditional machine learning algorithms have been presented in the literature. These research presented solutions based on k-NN, SVM, KHNN, and Artificial neural networks (ANN), among other technologies[11].

2. Hybrid Texture and Colour Descriptor

This section describes several strategies for enhancing medical pictures, including image enhancement and noise reduction techniques that improve image quality for higher retrieval accuracy. This first step is critical in detecting, tracing, and extracting the area from the globe. Because photos are finer, sharper, and improved in this stage. For the given application, the improved picture is finer than the original and provides more precise segmentation. The primary goals of this stage are to increase image and quality[12].

The segmentation process is mostly determined by the abrupt change of picture intensity level. A blurry or noisy picture is not suitable for extracting information. The average filter was used in this approach to smooth the pictures by lowering the fluctuation in image intensity values from one pixel to the next. An average filter is a linear smoothing filter that was created by replacing the value of each pixel in a picture with the grey level average using a filter mask. The suggested technique employed a 5x5 filter mask for filtering, which improved picture quality and decreased noise. The convolution sum of the filter mask with matching intensity values in an image was used to perform this filter operation[13-14].

A pinnacle finding Step is conducted on the collected cross-area profiles. Our goal is to determine if a pinnacle is accessible at the profile's focal point, i.e., in the vicinity of the competitive point for a certain heading. A few pinnacle attributes are selected, and the final list of capabilities includes a number of factual estimations that indicate how these qualities change as the introduction of the cross-segment evolves. Along these lines, a range of key properties, such as symmetry and structural state, as well as its difference from the foundation, might be

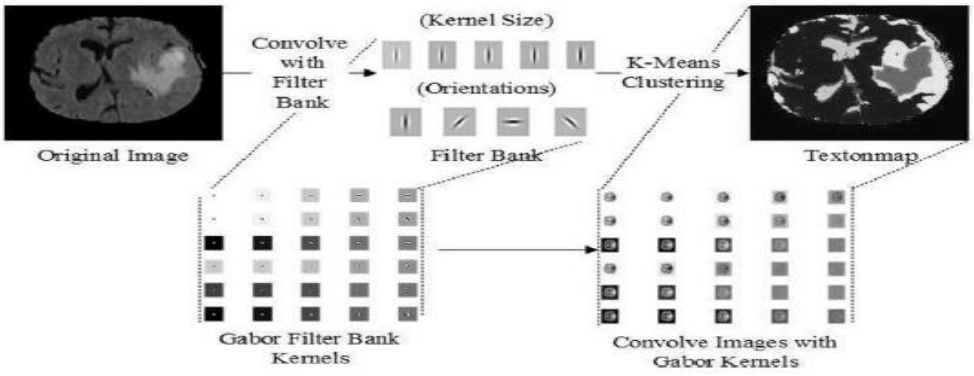
statistically communicated[15]. Gabor filters are often employed to gather local spatial and frequency data. This filter may be used in a variety of image processing applications such as texture analysis, edge extraction, object identification, and so on. A 2-D Gabor function is a Gaussian modulated by a complex sinusoid, and its Fourier transform is given by Eqn(2):

$$g(x,y) = \frac{1}{2\pi\sigma_x\sigma_y}.e^{-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)}.e^{j2\pi fx} \quad (1)$$

$$G(u,v) = e^{-\frac{1}{2}\left(\frac{(u-W)^2}{\sigma_u^2}+\frac{v^2}{\sigma_v^2}\right)} \quad (2)$$

The orientation picture creation is shown in below Figure 1, and then the Gabor transform images are separated into blocks of size 2x2, 4x4, and 8X8, and the number of blocks of each intensity value 1 to 255 is calculated.

Fig. 1. Construction of the orientation image using Gabor filter bank.



Above Figure 1 shows the feature extraction technique is an essential method for improving classification accuracy. It collects the image's important information and converts it into a feature vector. The acquired feature vector is then used in the retrieval or classification procedure. In this article, three kinds of characteristics are derived from the query picture based on their form, margin, and density. The form features define the image's bounds; the margin features explain the image's boundary characteristics; and the density feature, which depicts the image's brightness fluctuation. Finally, the three collected features are combined into a single feature vector[16]. The density degree of an MRI brain picture is defined by the brightness variation of the MRI image. MRI images are displayed in grayscale, with the density degree signified by the brightness fluctuation of the MRI picture. The following steps are used to get density features: To begin, split the MRI picture into two sections: inner and outside regions. The inner region's minor axis is equal to half of the outer region's minor axis. Using Eqn, calculate the density degree (3).

$$Density\ Degree = \frac{\Psi_{inner}}{\Psi_{outer}} \quad (3)$$

Most of the time, successful classification is subjective to picture attributes related with edges and network depth. As a result, the pre-processing approaches indicated above are applied to the input picture sequences in order to maintain the features edge and the connection of each edge for improved categorization. We boosted our network in every manner conceivable, thus the term Maximum Boosted Convolutional Neural Network.

Illness classification is an important part of disease categorization using image processing methods. The classification of illnesses based on pathogen groups is a key study subject and possibly a difficult area of work. Classification and detection are extremely similar, however the major emphasis in classification is on categorising

distinct illnesses, followed by classification according to various pathogen groupings. It is divided into two sections: training and testing. Most of the time, successful classification is subjective to picture attributes related with edges and network depth. As a result, the pre-processing approaches discussed above are applied to the incoming picture sequences in order to maintain the features edge and the connectivity of each edge for improved classification[17-26]. We boosted our network in every manner conceivable, thus the term Maximum Boosted Convolutional Neural Network. A typical CNN structure consists of a convolutional layer, a max-pooling layer with an activation function, and a fully connected layer. The appropriate feature map h(p, q) will be produced by convolving the input data with a convolution kernel v(p, q) of size m n, as stated in Eqn (4).

$$h(p, q) = u(p, q) * v(p, q) \quad (4)$$

The input to the hidden layers is modified after each convolution. Such input distributions restrict the performance of the layer parameters and diminish the learning rate during training. After each convolutional layer that downsamples the input pictures, maxpooling layers are added. The features are retrieved and passed to the fully linked layer using the suggested architecture. The input from the fully connected layer is fused and merged with the LSTM layer for classification. Maxpooling and an activation function are employed in each convolutional layer to increase the network's ability to extract deterministic features.

3. Experimental results

The performance of our proposed system is assessed using four benchmark Corel datasets (Corel-1000 and Corel-10000), the Oliva dataset, and the Caltech-256 dataset. Varying classes have different picture sizes, dimensions, and resolutions. The performance of CBIR on medical pictures was assessed using images from the National Cancer Institute database (http://cancerimagingarchive.net). The dataset in this case is made up of 20 individual patients and 200 MRI pictures. For each patient, three MR image orders were considered: T1, T2, and FLAIR. However, each volume has a different number of slices, ranging from 100 to 150.

3.1 Retrieval metrics

The two most important measures in information retrieval are accuracy and recall. As the subjective choral suggest, specifically F — measure, and it is an overall performance measure, the two metrics are often mixed.

$$F = \frac{(1+\beta^2) * R * P}{(\beta^2 * P) + R}$$

$$P = \frac{I_N}{N}$$

$$R = \frac{I_N}{M}$$

Eqn can be used to define it (4) (5)

3.3 Similarity measure

It is well known that the exactness not just depends on storing features illustration, but also superior parallel measure or detachment metric. Feature vector of query image Q is representedas $f_Q = f_{Q_1} + f_{Q_2} + \ldots\ldots f_{Q_{Lg}}$ obtained after the attribute removal.Similarity both image in the database is represented with feature vector $f_{DB_j} = (f_{DB_{j1}} + f_{DB_{j2}} + \ldots\ldots f_{DB_{jLg}}); j = 1, 2, \ldots\ldots |DB|$. In our work, four types of similarity distance metrics (Manhattan, Euclidean, Camberra detachment assess and Detachment assess) are used and these are per Eqn(5)

$$D(Q,DB)=\sum_{i=1}^{Lg}\left|f_{DB_{ji}}-f_{Q,i}\right|$$

L1 - (6)

$$D(Q,DB)=\left(\sum_{i=1}^{Lg}\left(f_{DB_{ji}}-f_{Q,i}\right)^{2}\right)^{1/2}$$

L2-

$$D(Q,DB)=\sum_{i=1}^{Lg}\frac{\left|f_{DB_{ji}}-f_{Q,i}\right|}{\left|f_{DB_{ji}}\right|+\left|f_{Q,i}\right|}$$

d1 -- (7)

$$D(Q,DB)=\sum_{i=1}^{Lg}\left|\frac{f_{DB_{ji}}-f_{Q,i}}{1+f_{DB_{ji}}-f_{Q,i}}\right|$$

d2-

Where $f_{DB_{ji}}$ is the ith feature of jth image in the record

Table 1 displays the experimental findings of similarity measures such as L1-Manhatten distance, L2-Euclidian distance, Cambera distance, and d1 distance measure of four distinct information sets. Figures 4 and 5 show the related evaluation graph.

Table 1. The retrieval result with four methods on four datasets

Data set name	**Similarity measure(%)**	**METHODS**			
		L1	**L2**	**Cambera**	**D1**
Corel-1000	Precision	78.67	74.67	78.00	74.67
	Recall	8.56	9.84	8.48	9.84
Oliva	Precision	53.69	51.80	52.85	48.85
	Recall	2.94	2.87	2.91	2.84
New Caltech	Precision	46.93	45.24	46.06	45.27
	Recall	4.08	3.97	4.01	3.97
Corel-10,000	Precision	48.15	47.11	49.96	47.28
	Recall	6.90	6.53	6.88	6.55

We compared our proposed image retrival approach to many descriptor-based image retrival methods, including CED, CMS, MTH, and MSD. The performance study was carried out by generating graphs of assessment measures including accuracy and recall is shown as below figure 2 and 3.

Fig 2. Similarity measures of Four data set using Precision

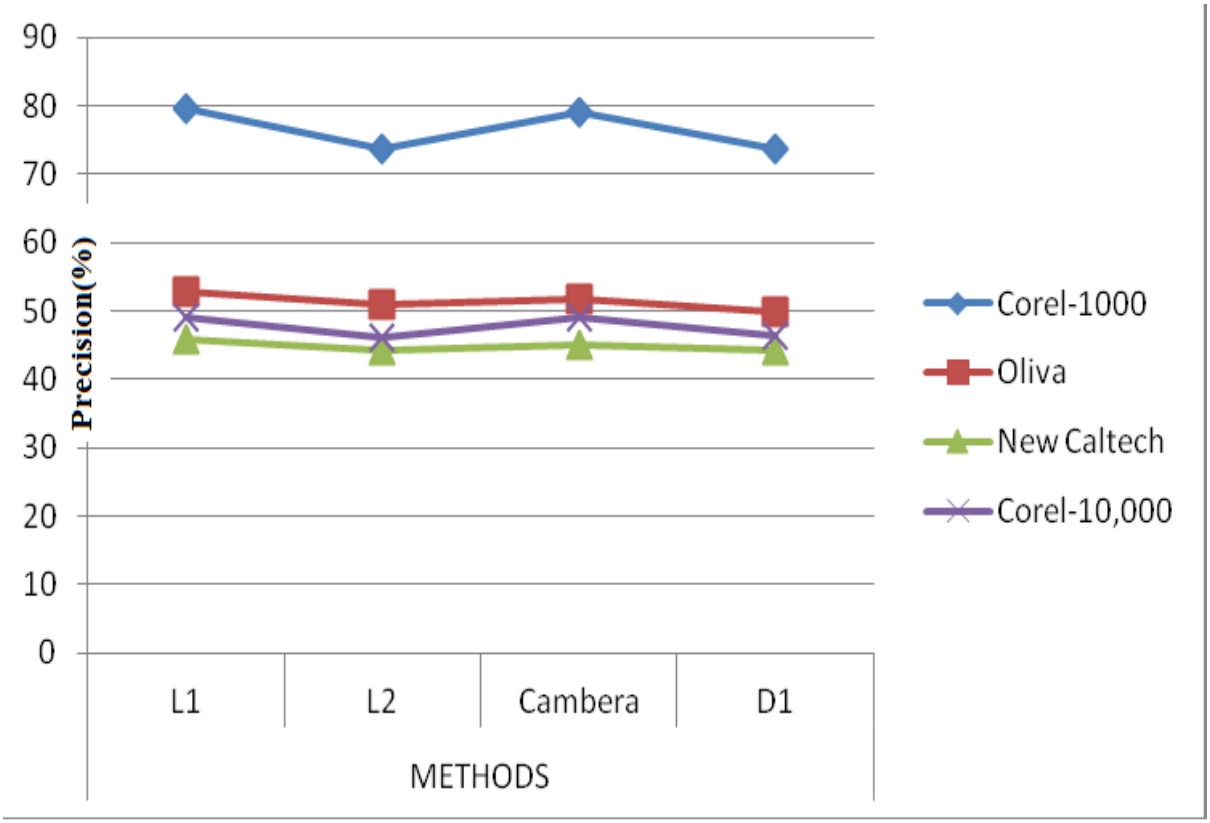
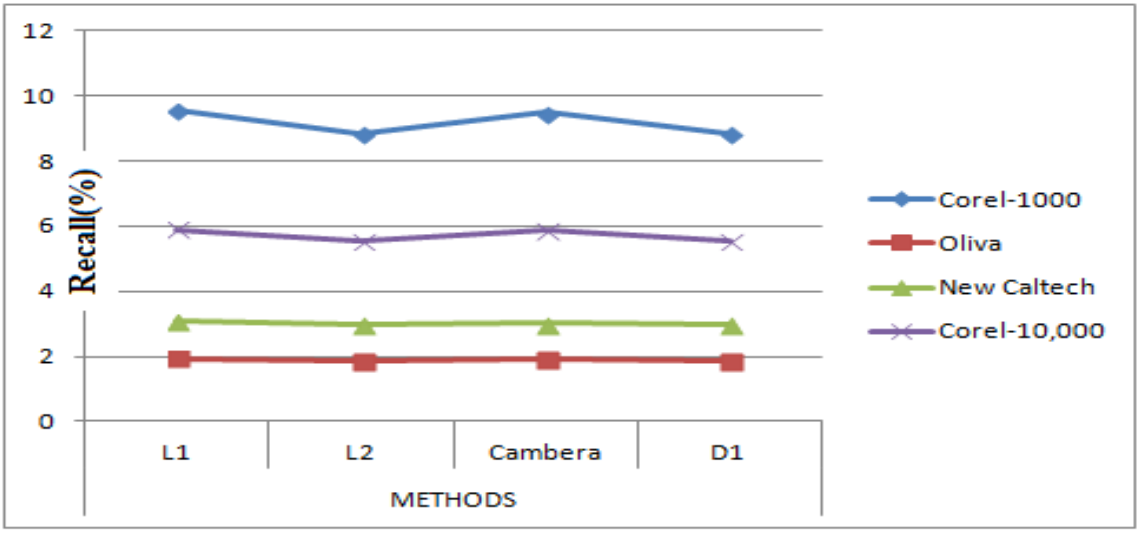


Fig 3. Similarity measures of Four data set using Recall



3.6 Experimental results using Medical Image Data set

Medical images and communication (DICOM) was created by the National Electrical Manufacturers Association (NEMA) to help in the viewing and dissemination of medical pictures such as MRI, CT, and ultrasound for research and storage purposes.

Fig. 4: Retrivel results of NEMA dataset. (a) Query image, (b) retrieved images.

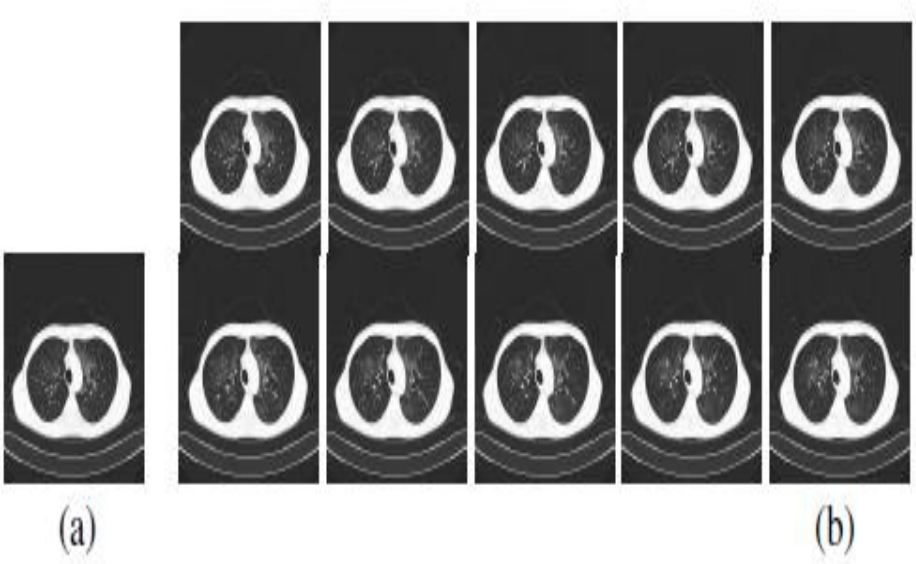
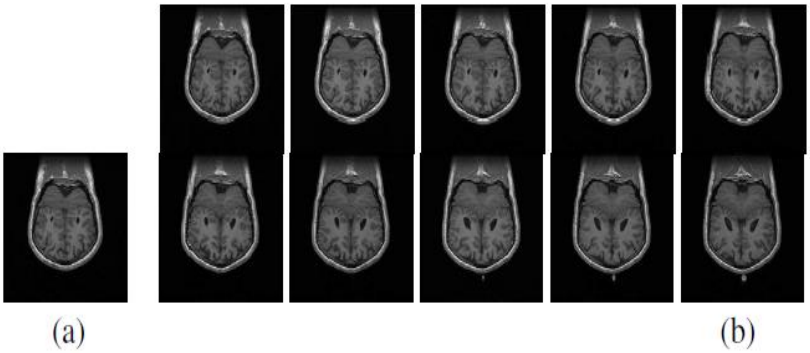


Fig. 5 : Retrivel results of OASIS dataset. (a) Query image, (b) retrieved images.



We used ten sets of computed tomography (CT) medical pictures produced by NEMA and the Open Access Series of Imaging Studies in this investigation (OASIS). The database details are accessible in [27-30], the query image from NEMA using the suggested technique is shown in Fig 4, and OASIS is shown in Fig 5.

Conclusion

The advancement of computer-aided detection methods in recent years has made them a nondestructive and popular tool for cancer detection in MRI images. The approach comprises many image processing steps, beginning with pre-processing the original pictures using an anisiotropic filter for noise reduction. Following that, an efficient picture segmentation method based on region growth is utilised to separate the cancer area from the backdrop. Then, many characteristics are extracted to improve the classification accuracy process. To accomplish optimum feature extraction, an ideal strategy for picking helpful features and trimming the remaining worthless features is utilised. Based on the experimental findings, the suggested technique is effective for retrieving photos from several datasets such as Corel-1000, Oliva, New Caltech, and Corel-10,000.

References

[1] Merriam-Webster, Inc.. Merriam-Webster's Dictionary of English Usage. (MerriamWebster, Inc.).

[2] William Morris, Ed.. The American Heritage Dictionary of the English Language.(William Morris, Ed.).

[3] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk. 2009. Frequency-tuned salient region detection. In IEEE CVPR '09.

[4] P. Adamopoulos and A. Tuzhilin. 2015. On Unexpectedness in Recommender Systems: Or How to Better Expect the Unexpected. ACM TIST (2015).

[5] Erik Boiy and Marie-Francine Moens. 2009. A machine learning approach to sentiment analysis in multilingual Web texts. Information retrieval (2009).

[6] D. Borth, T. Chen, R. Ji, and S.F. Chang. 2013. Sentibank: large-scale ontology and classifiers for detecting sentiment and emotions in visual content. In MM'13.

[7] G. Bradski. 2000. Dr. Dobb's Journal of Software Tools (2000).

[8] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell. 2014.DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition.In ICML '14.

[9] Carsten Eickhoff and Arjen P de Vries. 2013. Increasing cheat robustness of crowdsourcing tasks. Information retrieval (2013).

[10] M. Ge, C. Delgado-Battenfeld, and D. Jannach. 2010. Beyond Accuracy: Evaluating Recommender Systems by Coverage and Serendipity. In RecSys '10.

[11] R. Girshick, J. Donahue, T. Darrell, and J. Malik. 2014. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In CVPR '14.

[12] Robert J. Goldwater. 1986. Primitivism in modern art. Harvard University Press.

[13] Jorge Gracia, Elena Montiel-Ponsoda, Philipp Cimiano, Asunción Gómez-Pérez, Paul Buitelaar, and John McCrae. 2012. Challenges for the multilingual web of data. Web Semantics: Science, Services and Agents on the World Wide Web (2012).

[14] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. 2009. The WEKA Data Mining Software: An Update. SIGKDD Explor. Newsl. (2009).

[15] Mark A Hall. 1999. Correlation-based feature selection for machine learning. Ph.D. Dissertation. The University of Waikato.

[16] Robert M Haralick. 1979. Statistical and structural approaches to texture. Proc. IEEE 67, 5 (1979).

[17] Peter Howarth and Stefan Rüger. 2004. Evaluation of Texture Features for ContentBased Image Retrieval. In CIVR '04.

[18] David A. Hull and Gregory Grefenstette. 1996. Querying Across Languages: A Dictionary-based Approach to Multilingual Information Retrieval. In SIGIR '96.

[19] Michael Jacobs. 1995. The painted voyage: art, travel and exploration 1564-1875.British Museum Press London.

[20] Olivia H Jenkins. 1999. Understanding and measuring tourist destination images. International Journal of Tourism Research (1999).

[21] Yangqing Jia, Evan Shelhamer, Jeff Donahue, Sergey Karayev, Jonathan Long, Ross Girshick, Sergio Guadarrama, and Trevor Darrell. 2014. Caffe: Convolutional Architecture for Fast Feature Embedding. In ACM MM '14.

[22] Jonathan Hare, Sina Samangooei, and David Dupplaw. 2011. OpenIMAJ and ImageTerrier: Java Libraries and Tools for Scalable Multimedia Analysis and Indexing of Images. In MM '11.

[23] Alden Jones. 2007. This is Not a Cruise. http://archive.fo/TEec. (2007).

[24] B. Jou, T. Chen, N. Pappas, M. Redi, M. Topkara, and S.F. Chang. 2015. Visual Affect Around the World: A Large-scale Multilingual Visual Sentiment Ontology. In MM'15.

[25] Marius Kaminskas and Derek Bridge. 2017. Diversity, Serendipity, Novelty, and Coverage: A Survey and Empirical Analysis of Beyond-Accuracy Objectives in Recommender Systems. ACM Trans. Interact. Intell. Syst. (2017).

[26] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. 2012. ImageNet Classification with Deep Convolutional Neural Networks. In NIPS '12.

[27]David, D. S., Arun, S., Sivaprakash, S., Raja, P. V., Sharma, D. K. et al. (2022). Enhanced Detection of Glaucoma on Ensemble Convolutional Neural Network for Clinical Informatics. CMC-Computers, Materials & Continua, 70(2), 2563–2579.

[28]David, D. S., Anam, M., Kaliappan, C., Arun, S., Sharma, D. K. et al. (2022). Cloud Security Service for Identifying Unauthorized User Behaviour. CMC-Computers, Materials & Continua, 70(2), 2581–2600.

[29]Thendral, R., & David, D. S. (2022). An Enhanced Computer Vision Algorithm for Apple Fruit Yield Estimation in an Orchard. In Artificial Intelligence and Technologies (pp. 263-273). Springer, Singapore.

[30]J. K. S and D. S. David, "A Novel Based 3D Facial Expression Detection Using Recurrent Neural Network," 2020 International Conference on System, Computation, Automation and Networking (ICSCAN), Pondicherry, India, 2020, pp. 1-6, doi: 10.1109/ICSCAN49426.2020.9262287