Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(4), 2722-2732 | ISSN:2251-6727

A Conceptual Research Basedon Image Classification Based On Neural Network Architecture

[1]Suman Singh, [2]Dr. Nidhi Mishra

[1,2] Kalinga University Raipur, C G, India

(Received: 02 September 2023 Revised: 14 October Accepted: 07 November)

KEYWORDS
Deep CNN,
Features fusion,
Features reduction,
Classification.

ABSTRACT:
we propose a novel technique leveraging deep convolutional neural networks (DCNNs) in conjunction with scale-invariant feature transform (SIFT). Initially, an enhanced saliency method is employed to identify key points, from which point features are extracted. Subsequently, features are extracted from two deep CNN models—VGG and AlexNet. Following this, entropy-controlled method is applied to the DCNN pooling and SIFT point matrix to discern robust features.
The identified robust features are then fused into a matrix using a sequential approach, which is subsequently inputted into an ensemble classifier for recognition. This approach amalgamates the strengths of DCNNs and SIFT, leveraging their complementary capabilities to enhance classification accuracy, particularly in scenarios characterized by complex backgrounds, congestion, and object similarity.

1 INTRODUCTION

1.1 Structure of Neural Networks

The construction of a standard brain network comprises of various counterfeit neurons, which are very much associated direct informing or trade of data. While preparing the model, the loads related with these associations are adapted to streamlined characterization precision.

The layers can be concealed for different activities to limit model intricacy. Each secret layer is made out of a few neurons that are interlinked to the past layer and can identify the picture descriptors.

Fig.1.1 shows the association of layers (input, stowed away, and yield) in an ordinary brain networks model. The underlying arrangement of essential examples is caught by the information layer. Secret layers do additionally handling to get additional theoretical examples from the underlying set. The result layer further consolidates the examples to introduce significant level last dynamic elements. A standard profound brain network model can contain in excess of 10 layers and up to two or three hundred layers.

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(4), 2722-2732 | ISSN:2251-6727

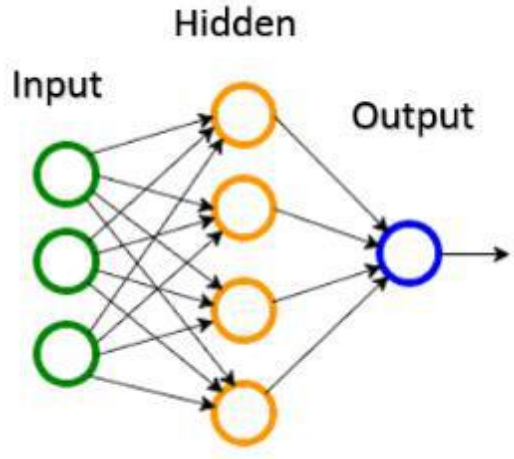

Fig. 1.1 Layerd design of Brain organizations

2 METHODOLOGY

The methodology for the research on image classification in light of profound convolutional brain organizations can be isolated into three main steps: data collection and preprocessing, architecture design of the proposed deep CNN model, and training and testing procedures flow process.

1. Data collection and preprocessing: The initial step includes gathering a huge dataset of pictures for preparing and testing the profound CNN model. The pictures should be diverse and representative of the target task. The collected data must be pre-processed to ensure that the images are of consistent size, color depth, and orientation.

2. Architecture of the proposed deep CNN model: The second step involves designing the engineering of the proposed profound CNN model. The architecture should with increasing depth and decreasing spatial resolution, followed by pooling layers to diminish the spatial dimensions. Additionally, the architecture can include skip connections, batch normalization, and dropout to work on the preparation and speculation of the model.

3. Training and testing procedures: The third step involves training the proposed deep CNN model on the pre-processed dataset using stochastic gradient descent with backpropagation. The model should be evaluated measure its exactness and speculation performance. Additionally, hyper parameter tuning can be performed to optimize the performance of the model. Finally, the trained model can be used for real-world applications like item acknowledgment, face acknowledgment, and clinical picture investigation.

In summary, the methodology for the research on picture grouping in light of profound CNNs involves collecting and pre-processing a diverse dataset, planning the design of the deep CNN model.

3 RESULT DISCUSSION

3.1 Materials and Methods

In this assessment, we used three famous datasets like Caltech101, PASCAL 3D, and 3D dataset to oversee complex thing revelation and gathering. These datasets contain many thing classes and enormous number of pictures. To overcome the troubles of these datasets like brightening, variety, and closeness among different item classes, we propose another strategy for object arrangement in light of DCNN highlights extraction alongside Filter focuses. The proposed strategy comprises of two significant stages, which are executed in equivalent. In the underlying step, Channel point features are taken out from arranged RGB divided objects. In second step, DCNN features are isolated through pre-

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(4), 2722-2732 | ISSN:2251-6727

arranged CNN models like AlexNet and VGG. The both Channel point and DCNN features are joined into one structure by an equal combination strategy and the best elements are chosen for conclusive grouping. The point by point portrayal of each step is given beneath in area 3.2 to 3.4. The extensive stream outline is introduced in Fig. 3.1.

3.2 Improved Saliency Method

A superior saliency strategy is used by utilizing existing saliency approach name HDCT, for single thing area. In this step, we eliminate a lone thing from an image by an ongoing saliency methodology explicitly HDCT saliency evaluation. The idea behind the

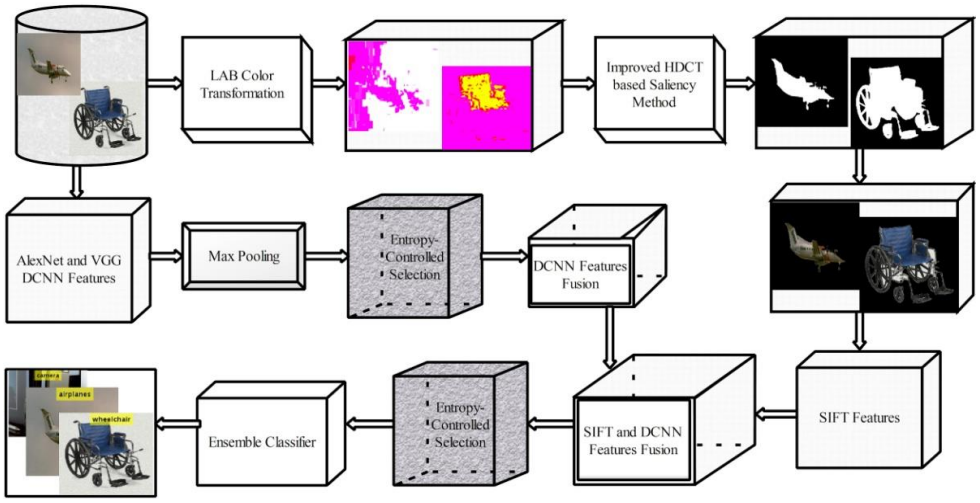

Fig. 3.1 Flow diagram of the proposed object classification method

improvement of saliency procedure is to complete the assortment spaces before it gives the data picture to the saliency method. The LAB assortment change is utilized hence, which perceives assortment in 3 angles containing L* for delicacy, a*, and b* are utilized for assortment parts green-red and blue-yellow independently. The parts L* is more unbelievable white at 100 and hazier dim at 0, however 'a' and 'b' channels show the typical characteristics for the RGB picture.

This change is characterized as follows:

Let U(i, j) implies a data RGB picture having length N ×M, then, for RGB to LAB change, first RGB to XYZ change is performed through Eqs. 3.1–3.10:

$$\begin{bmatrix} \varphi(X) \\ \varphi(Y) \\ \varphi(Z) \end{bmatrix} = [M \times N] \begin{bmatrix} \varphi^r \\ \varphi^g \\ \varphi^b \end{bmatrix} \quad (3.1)$$

where φ(X), φ(Y), and φ(Z) mean the X, Y, and Z channels, which are removed from red (φr), green (φg), and blue channel (φb). The φr, φg, and φb channels are portrayed as:

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(4), 2722-2732 | ISSN:2251-6727

$$\varphi^{r} = \sum_{k=1} \frac{\varphi k}{\Delta_k}, k = \text{Red} \tag{3.2}$$

$$\varphi^{g} = \sum_{k=2} \frac{\varphi k}{\Delta_k}, k = Green \tag{3.3}$$

$$\varphi^{b} = \sum_{k=3} \frac{\varphi k}{\Delta_k}, k = Blue \tag{3.4}$$

Then, at that point, LAB transformation is characterized as:

$$\left(\varphi^{L} = \beta_1 \times \left(f_y - 16\right)\right), \beta_1 = 116 \tag{3.5}$$

$$\left(\varphi^{*A} = \beta_2 \left(f_x - f_y\right)\right), \beta_2 = 500 \tag{3.6}$$

$$\left(\varphi^{*B} = \beta_3 \left(f_y - f_z\right)\right), \beta_3 = 200 \tag{3.7}$$

Where, f_x, f_y, and f_z are straight capabilities which are registered as:

$$f_x = \left\{ \sqrt[3]{x_r} \middle| \frac{kx_r + 16}{116}, \rightarrow x_r > \in \middle| \, otherwise \right\}, x_r = \frac{X}{Xr} \tag{3.8}$$

$$f_y = \left\{ \sqrt[3]{y_r} \middle| \frac{ky_r + 16}{116}, \rightarrow y_r > \in \middle| \, otherwise \right\}, y_r = \frac{Y}{Yr} \tag{3.9}$$

$$f_z = \left\{ \sqrt[3]{z_r} \middle| \frac{kz_r + 16}{116}, \rightarrow z_r > \in \middle| \, otherwise \right\}, z_r = \frac{Z}{Zr} \tag{3.10}$$

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(4), 2722-2732 | ISSN:2251-6727

3.3 Sift Features

Scale Invariant Element Change (Filter) is initially planned in 2004 by [43] and have appeared as a strength descriptors for object distinguishing proof and affirmation. The Filter highlights are processed in four stages. In the initial step, neighborhood central issues are resolved that are significant and stable for given pictures. Then includes are removed from each central issue that makes sense of the neighborhood picture locale tests, which are connected with its scale space coordinate picture. In the subsequent step, powerless elements are taken out by a particular limit esteem. In the third step, directions are doled out to each central issue in view of nearby picture angle bearings. At last, the 1 × 128 layered include vector is extricated, and bi-straight introduction is performed to work on the strength of elements. The above hypothesis is characterized through Eqs. 3.11–3.13:

$$\xi(\mu,\nu,\sigma) = \psi_G(\mu,\nu,\sigma) \otimes S_{final}(X_i) \tag{3.11}$$

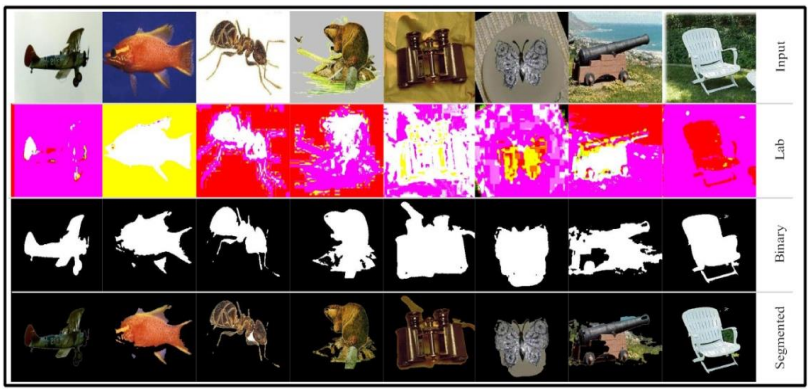

Fig. 3.2 Proposed better saliency technique results

$$\psi_G(\mu,\nu,\sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2}\left(\frac{\mu^2+\nu^2}{2\sigma^2}\right)} \tag{3.12}$$

$$D(\mu,\nu,\sigma) = (\psi_G(\mu,\nu,k\sigma) - \psi_G(\mu,\nu,\sigma)) \otimes S_{final}(X_i) = \xi(\mu,\nu,k\sigma) - \xi(\mu,\nu,\sigma) \tag{3.13}$$

where ξ(u, v, σ) is scale space of an image, ψG(u, v, kσ) connotes the variable-scale Gaussian, k is a multiplicative part and D(u, v σ) implies the differentiation of Gaussian convolved with a separated picture.

3.4 Deep CNN Features

As of late, in the area of PC vision, machine, and example acknowledgment, profound learning have shows further developed execution for picture arrangement on huge

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(4), 2722-2732 | ISSN:2251-6727

datasets [20]. The profound learning plans like profound CNN and repetitive NN have been utilized to human activity acknowledgment, discourse acknowledgment, record arrangement, agrarian plants, clinical imaging, and numerous different regions and shows unrivaled execution. In object arrangement, CNN shows a lot of consideration because of their capacity to decide suitable relevant elements in picture classification issues naturally. A basic CNN model comprises of four sorts of layers. At first, an information picture is passed and registers its neurons by convolution layer, which are associated with neighborhood districts of the information. Every neuron is registered by dab item between their little districts and loads, which are associated with in the info volume. From there on, actuation is performed utilizing ReLu layer. The ReLu layer never changes the size of an info picture. Then, pooling layer is performed to diminish the uproar influences in the removed features. Finally, critical level not entirely settled by a totally related (FC) layer.

In this article, we use two pre-arranged significant CNN models, for instance, VGG19 and AlexNet, which are used for features extraction. These models merge convolution layer, pooling layer, normalization layer, ReLu layer, and FC layer. As inspected over the convolution layer eliminates close by features from an image, which is sorted out by Eq. 3.14:

$$g_i^{(L)} = b_i^{(L)} + \sum_{j=1}^{m_1^{(L-1)}} \psi_{i,j}^{(L)} \times h_j^{(L-1)} \tag{3.14}$$

Where $g_i^{(L)}$ means the result layer $L, b_i^{(L)}$ is base value $\psi_{i,j}^{(L)}$ signifies the channel interfacing the jth highlight map, and h_j indicates the L − 1 result layer. Then, pooling layer is described which eliminate most outrageous responses from the lower convolutional layer with an objective of diminishing immaterial components. The maximum pooling additionally settle the issue of overfitting and generally 2 × 2 reviewing is performed on the eliminated system. Mathematically, max pooling is depicted through Eqs. 3.15–3.17:

$$m_1^{(L)} = m_1^{(L-1)} \tag{3.15}$$

$$m_2^{(L)} = \frac{m_2^{(L-1)} - F(L)}{S^L} + 1 \tag{3.16}$$

$$m_3^{(L)} = \frac{m_3^{(L-1)} - F(L)}{S^L} + 1 \tag{3.17}$$

Where S^L denotes the stride $, m_1^{(L)}, m_2^{(L)}, \text{and } m_3^{(L)}$ are characterized channels for highlight guide like 2 × 2, 3 × 3. Different layers like ReLu and completely associated (FC) are characterized as:

$$\mathrm{Re}_i^{(l)} = \max\left(h, h_i^{(l-1)}\right) \tag{3.18}$$

$$Fc_i^{(l)} = f\left(z_i^{(l)}\right) \text{ with } z_i^{(l)} = \sum_{j=1}^{m_1^{(l-1)}} \sum_{r=1}^{m_2^{(l-1)}} \sum_{s=1}^{m_3^{(l-1)}} w_{i,j,r,s}^{(l)} \left(Fc_i^{(l-1)}\right)_{r,s} \tag{3.19}$$

Where $Re_i^{(l)}$ denotes the ReLu layer, $Fc_i^{(l)}$ implies the FC layer. The FC layer follows the convolution and pooling layers. The FC layer resembles convolution layer and most of the researchers perform incitation on the FC layer for significant component extraction.

3.5 Pre-Prepared Profound CNN Organizations

In this investigation, we use two pre-arranged significant CNN models like VGG and AlexNet for significant components extraction. AlexNet significant CNN model is arranged by Krizhevsky et al. [27] using ImageNet dataset. This association contains five convolution layers, three pooling layers, and 3 FC layers close by softmax portrayal capacity. This association arranged on input picture size 227x227x3.

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(4), 2722-2732 | ISSN:2251-6727

VGG-19 CNN network is proposed by Zisserman et al. [20] which contains 16 convolution layers, 19 learnable burdens layers, 3 FC layers close by softmax capacity. This association is ready on ImageNet dataset and shows unprecedented execution. This association in like manner includes dropout regularization in the FC layer and apply ReLu authorization ability on all the convolution layers. The size of the planning input pictures is picked as 224x224x3.

3.6 Features Extraction and Fusion

In this fragment, we present our proposed feature extraction and mix strategy. The features are taken out from pre-arranged significant CNN models using the different number of layers. In this work, two pre-arranged models are used like VGG19 and AlexNet for features extraction. The critical place of significant CNN features extraction from two models is to additionally foster the request accuracy. Since each model has obvious characteristics and gives different components. Thusly, by using this advantage, we remove features by performing inception on the FC7 layer and applying max pooling to dispose of the racket factors. From that point on, an entropy-controlled procedure is completed for best part decline. The proposed feature extraction and abatement designing are shown in Fig. 3.3. As shown in Fig. 3.3, three kinds of features are isolated like AlexNet significant CNN, VGG19 CNN, and Channel. For AlexNet and VGG19, convolution layer is used as a data layer. Then, at that point, commencement is performed on FC7 layer for the two associations to remove significant CNN features. The size of significant CNN features for yield layer FC7 is 1 × 4096 for the two associations. The component size of both outcome layer is higher. In this manner, we perform maxpooling of channel size of 2 × 2, which kills the uproar influences and picks the most outrageous worth part of the given channel.

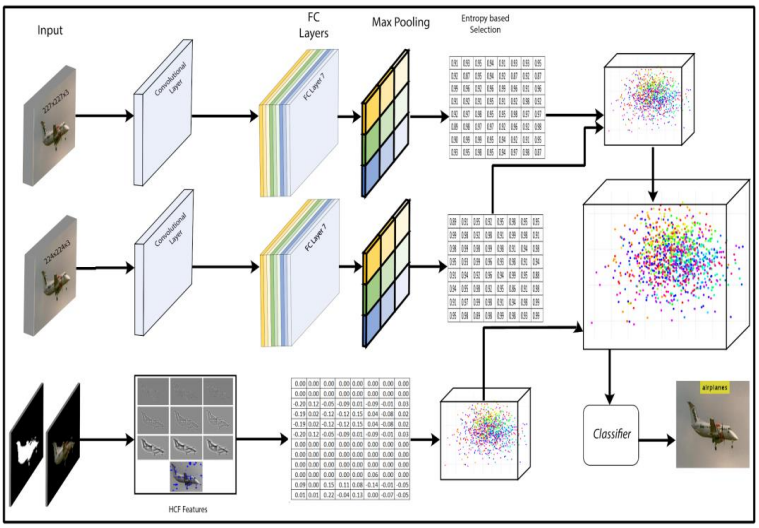

Fig. 3.3 Proposed profound CNN and Filter highlights combination and decrease strategy for object grouping

After max-pooling, the new component vectors of size 1 × 2048 are obtained, which are also improved by entropy-controlled incorporate lessening method. As eliminated feature vectors can make further developed results, but they increase the execution time. Along these lines, our middle is to additionally foster the portrayal accuracy and reducing the execution time. This issue is settled by an entropy controlled method. The entropy signs to the data about mediation in by showing the system issue [50]. In light of its capacity to portray system lead, entropy gives

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(4), 2722-2732 | ISSN:2251-6727

the significant information which can be used in features plan [7]. Among a couple, we use the Renyi entropy strategy for incorporate reduction. In the states of fractal perspective evaluation, the Reyni entropy technique chooses the reason of the speculation of summarized angles. The fractal perspective checks the change instances of the given part space. The Reyni entropy is portrayed as follows:

Let f_1, f_2, … , fn connote the A component space after max-pooling, g_1, g_2, g_3, … gn mean the B feature space aftermax-pooling, and ξ_1, ξ_2,… , ξ_n show the ξ incorporate space, where A∈ AlexNet DCNN features, B∈ VGG19 DCNN features, and ξ∈ Channel point incorporate vector. The part of every components space is 1 × 2048, 1 × 2048, and 1 × 128. The entropy is sorted out by Eq. 5.28:

$$E_\alpha(X) = \frac{1}{1-\alpha}\log\left(\sum_{i=1}^{n} p_i^a\right) \tag{3.20}$$

Where $\alpha \geq 0 \,\&\, <1,\ X \in (f_n, g_n, \xi_n),$ what's more, pi mean the likelihood worth of separated highlight space A, B, and ξ which is described by pi = Pr(X = I) and addresses the length of all part spaces. The entropy ability gives another N ×M incorporate vector, which controls the haphazardness of every component space. Then, every N ×M feature vector is organized into climbing demand and the super 1000 components are looked over An and B vectors and 100 components from the ξ vector. Mathematically, this cycle is depicted by Eq. 3.21:

$$E(A) = \Phi(f_n, \varrho), E(B) = \Phi(g_n, \varrho), E(\xi) = \Phi(\xi_n, \varrho) \tag{3.21}$$

where E(A) implies the entropy information of component space A, E(B) shows the entropy information of part space B, E(ξ) connotes the entropy information of component space ξ, Φ demonstrates orchestrating capacity, and ϱ demonstrates the climbing demand movement. From that point on, both E(A) and E(B) entropy information features are merged in one matrix by the direct successive based procedure, which returns a component vector of size 1 × 2000, which is also joined with Channel point feature by the consecutive based strategy as shown in the above Fig. 3.3 and underneath explanation given in Eqs. 3.22–3.23:

$$\prod(Fused) = (N \times 1000) + (N \times 1000) + (N \times 100) \tag{3.22}$$

$$\prod(Fused) = N \times f_i \tag{3.23}$$

The size of the last part vector is 1 × 2100, which is dealt with to bunch classifier for request. The outfit classifier is a regulated learning strategy, which necessities to preparing information for expectation. Outfit technique consolidates a few classifiers information to deliver a superior framework. The plan of the gathering strategy is given underneath.

3.7 Experimental Results

The proposed technique is upheld on three open datasets, for instance, Caltech 101, PASCAL 3D+, and Barkley3D dataset. The Caltech-101 [14] dataset includes outright 102 specific article classes of 9144 pictures. Each class includes around 31~800 pictures. In any case, this dataset contains both RGB and dim pictures, which is a huge issue of this dataset. It is since, in such a case that articles are seen by their assortment, then assortment features are not performed well on

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(4), 2722-2732 | ISSN:2251-6727

Fig. 3.4 Proposed marked order results for the 3D dataset and Caltech101 dataset

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(4), 2722-2732 | ISSN:2251-6727

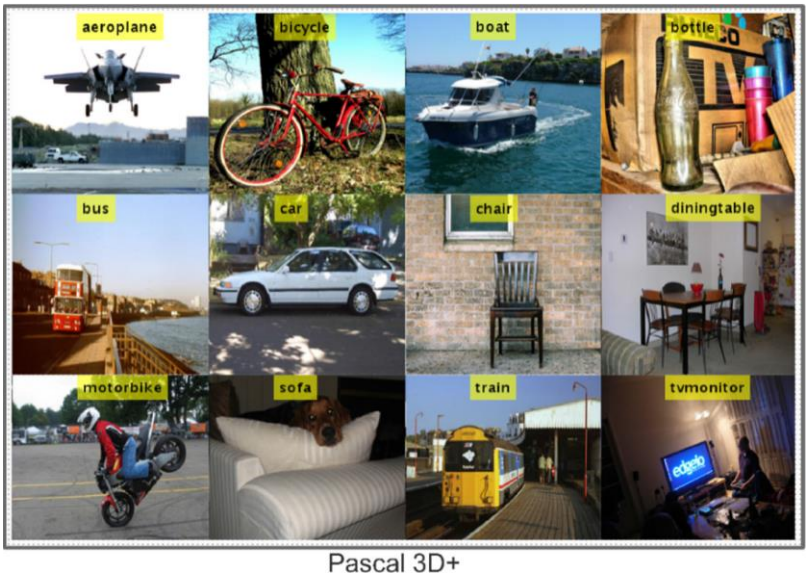

Fig. 3.5 Proposed checked request results for PASCAL 3D+ dataset

grayscale pictures. Pascal 3D+ dataset [11] is another trying informational collection which is used for object portrayal. This dataset is the blend of Pascal VOC 2012 and ImageNet. It contains an amount of 22,394 pictures of 12 exceptional classes. The classes which are ordinary between PASCAL VOC 2012 and ImageNet are united into another data base, called Pascal 3D+. Barkley3D object dataset [21] involves outright 6604 pictures of 10 thing classes including bicycle, vehicle, cellphone, head, iron, screen, mouse, shoe, stapler, and toaster. The amount of pictures in each class extent of 474-721. A short portrayal of each dataset is given in Table 3.1. For portrayal, we use Company helped tree (EBT) classifier and test its show with Straight SVM (LSVM), Quadratic SVM (QSVM), Cubic SVM (CSVM), Fine KNN (FKNN), Cubic KNN (CKNN), decision tree (DT), and weighted KNN (WKNN). The show of each and every not entirely settled by three measures including precision, sham negative rate (FNR), and execution time. All results are evaluated on 3.4 Gigahertz Corei7 seventh time PC with a Ram of 16 Gigabytes and a GPU of NVIDIA GeForce 1070 (8GB, 256 digit) having MATLAB 2017b.

4 CONCLUSION

The proposed procedure works in two equivalent advances. The greatest pooling is performed on removed features organizations to dispose of the boisterous information. From that point on, a Reyni entropy-controlled methodology is proposed which control the inconsistency of isolated incorporates and select the best components. The picked features are finally dealt with to bunch classifier for object course of action. The proposed method normally distinguishes and stamped object from endless model pictures with least human mediation. Moreover, we apply system on steady thing portrayal.

- Object identification and grouping is a difficult errand in PC vision with applications like visual surveillance, target recognition, face detection, etc.
- Existing methods struggle with complex backgrounds, occlusion, and similarity between objects.
- Deep learning, especially deep convolutional brain organizations (DCNNs), have shown promising outcomes for picture grouping errands.

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(4), 2722-2732 | ISSN:2251-6727

REFERENCES

[1] Abdullah-Al Nahid, ID and Yinan Kong, "Histopathological Breast-Image Classification Using Local and Frequency Domains by Convolutional Neural Network", Information 2018, 9, 19; doi:10.3390/info9010019.

[2] Yun Jiang, Li ChenI, Hai Zhang, Xiao Xiao, "Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module", PLOS ONE| https://doi.org/10.1371/journal.pone.0214587 March 29, 2019.

[3] Farhana Sultana, Abu Sufian, Paramartha Dutta, "Advancements in Image Classification using Convolutional Neural Network", arXiv:1905.03288v1 [cs.CV] 8 May 2019.

[4] Qing Li, Weidong Cai, Xiaogang Wangy, Yun Zhouz, David Dagan Feng and Mei Chen, "Medical Image Classification with Convolutional Neural Network".

[5] Iman Sajedian, Jeonghyun Kim and Junsuk Rho, "Finding the optical properties of plasmonic structures by image processing using a combination of convolutional neural networks and recurrent neural networks", Sajedian et al. Microsystems & Nanoengineering (2019) 5:27.

[6] Sara Aqab, Muhammad Usman Tariq, "Handwriting Recognition using Artificial Intelligence Neural Network and Image Processing", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 11, No. 7, 2020.

[7] Jaya Gupta, Sunil Pathak, Gireesh Kumar, "Bare Skin Image Classification using Convolution Neural Network", International Journal of Emerging Technology and Advanced Engineering.

[8] Wei Wang, Yujing Yang, Xin Wang,* Weizheng Wang, and Ji Li, "Development of convolutional neural network and its application in image classification: a survey", Optical Engineering 58(4), 040901 (April 2019).

[9] José Naranjo-Torres, Marco Mora, Ruber Hernández-García, Ricardo J. Barrientos, Claudio Fredes and Andres Valenzuela, "A Review of Convolutional Neural Network Applied to Fruit Image Processing", Appl. Sci. 2020, 10, 3443; doi:10.3390/app10103443.

[10] Md. Anwar Hossain & Md. Mohon Ali, "Recognition of Handwritten Digit using Convolutional Neural Network (CNN)", Global Journal of Computer Science and Technology: D Neural & Artificial Intelligence, Volume 19 Issue 2 Version 1.0 Year 2019.

[11] Mingyuan Xin and Yong Wang, "Research on image classification model based on deep convolution neural network", Journal on Image and Video Processing (2019) 2019:40, https://doi.org/10.1186/s13640-019-0417-8.

[12] Zhan Wu, Min Pengl, Tong Chen, "Thermal Face Recognition Using Convolutional Neural Network", 978-1-5090-0880-3/16/$31.00 ©20 16 IEEE.

[13] Valeria Maeda-Gutiérrez, Carlos E. Galván-Tejada, Laura A. Zanella-Calzada, José M. Celaya-Padilla, Jorge I. Galván-Tejada, Hamurabi Gamboa-Rosales, Huizilopoztli Luna-García, Rafael Magallanes-Quintanar, Carlos A. Guerrero Méndez and Carlos A. Olvera-Olvera, "Comparison of Convolutional Neural Network Architectures for Classification of Tomato Plant Diseases", Appl. Sci. 2020, 10, 1245; doi:10.3390/app10041245.

[14] Dong-Yeol Yun, Seung-Kwon Seo, Umer Zahid and Chul-Jin Lee, "Deep Neural Network for Automatic Image Recognition of Engineering Diagrams", Appl. Sci. 2020, 10, 4005; doi:10.3390/app10114005.

[15] Weibo Liu, Zidong Wang, Xiaohui Liu, Nianyin Zeng, Yurong Liu,d and Fuad E. Alsaadi, "A Survey of Deep Neural Network Architectures and Their Applications".