Image Classification Using Optimized Convolution Neural Network

Ramesh Chegoni
Department of Information Technology
CMR Technical Campus
Hyderabad, India
ramesh.it@cmrtc.ac.in

V. SriLakshmi
Department of CSE(AI & ML)
B V Raju Institute of Technology
Narsapur, India
srilakshmi.v@bvrit.ac.in

R. Sai Venkata Ramana
Computer Science & Engineering(Cyber Security) Department
Madanapalle Institute of Technology & Science
onlinesvr4@gmail.com

Veeramreddy Rajasekhar
Department of Computer Science and Engineering
Malla Reddy Engineering College (Autonomous,) Hyderabad, India
rajasekhar.v86@gmail.com

Ravindra Changala
Department of Computer Science and Engineering
Guru Nanak Institutions Technical Campus, Hyderabad, India
changalaravindra@gmail.com

Venkata Nagaraju Thatha
Department of Information Technology
MLR Institute of Technology
Hyderabad, India
nagarajuthatha@gmail.com

***Abstract*—Conventional artificial neural networks and machine learning algorithms encountered difficulties in fulfilling the processing needs for extracting features from large-scale images and training models. Furthermore, they showed poor efficiency and accuracy in classifying images. In order to lay a firm groundwork and offer aid for the categorization and recognition of large datasets, this study introduces a deep learning model for classification of images. The feature extraction technique of CNNs is refined by mitigating noise and adjusting model parameters, resulting in the development of an advanced CNN architecture for deep learning in image classification. The efficacy of the proposed model is assessed through tests that scrutinize the correlation between iteration numbers and accuracy of various network models are commonly employed in image classification. Notably, the research model surpasses its counterparts in terms of classification accuracy, indicating its superiority in the realm of image classification. The study contributes valuable insights by enhancing the understanding of CNNs and offering a more effective model for image categorization, supported by rigorous evaluations**

Keywords- deep learning, Machine learning, image classification, neural networks

I. INTRODUCTION

Many traditional machine learning algorithms like support vector machines, multilayer perception machines, etc. rely on simplistic models that can only handle small datasets and computational resources. Performance and generalizability issues with complicated classification tasks become glaringly apparent when dealing with target objects that include rich meanings. The recently constructed convolution neural network (CNN) has greatly improved the accuracy of many machine learning tasks and is excellent at handling picture classification and recognition difficulties; as a result, it has seen extensive use in the domain of image processing. The model has matured into a robust and widely applicable deep learning system. [1]

A multilayer neural network, the most popular and traditional deep learning framework is the convolutional neural network (CNN). Newman et al. provide a convolutional neural network–based reconstruction approach. The CNN model with pre-training or fine-tuning, the hybrid method, and the other two approaches were covered by Wang et al. [2]. The first two sets of executive images are input to the network once, whereas the third set employs a feature extraction approach based on patches. The survey sheds light on the connection between SIFT and the CNN-based method, analyzes a diverse range of prior work, and marks a significant step forward in contemporary case retrieval. We examine a novel approach to general and special case retrieval after studying and comparing the retrieval performance of various categories on multiple datasets. When it comes to hyper spectral image classification, convolutional neural networks (CNNs) thrive, and they're also highly interested in machine learning.

Accurate image identification and classification necessitate the acquisition of raw images data using an image collection tool. Following this, the next crucial step involves evaluating the obtained image data and removing any extraneous details, laying the groundwork for the refinement of distinctive features. This preliminary processing phase ensures that the subsequent stages of analysis and classification are based on a focused and relevant set of image features, contributing to the overall accuracy and effectiveness of the identification and classification process. The use of a machine learning approach will finally allow us to achieve the main objective of the image classification. The usage of computer vision technology is on the rise across several industries as it enables individuals to efficiently perform image processing tasks such as feature extraction and noise reduction [3].

Machine learning is a major area of artificial intelligence. Despite machine learning's tremendous progress in the past half-century, numerous issues remain. Some examples of these are recommendation systems, language translation, and sophisticated picture recognition and understanding [4].

Deep learning is a relatively new but rapidly expanding area of machine learning. This approach to data processing

2024 Parul International Conference on Engineering and Technology (PICET) | 979-8-3503-6974-8/24/$31.00 ©2024 IEEE | DOI: 10.1109/PICET60765.2024.10716049

makes use of the hierarchical architecture of both real and synthetic neural networks. It acquires basic features and then use a feature combination approach to acquire more complex ones. This paves the way for the system to conduct picture categorization or regression. Deep learning utilizes a multilayer neural network to independently learn from images and extract their fundamental properties, unlike traditional machine learning methods. The process of deep learning entails training neural networks to identify images by examining their unique features. To do this, we build an all-encompassing feature extraction model that overcomes the problems with the old methods. The integrated model's extractor needs to be able to learn and extract the unique features from the training batch of images with precision. The feature descriptors retrieved from the image are categorized using a variety of approaches, including GIST, SIFT, and histogram of gradient oriented and Local Binary Patterns. By means of deep learning, a computer may learn to classify visuals, audio, and textual data. The computer learns from large picture datasets, then stores an internal representation of each image based on its pixel values. Patterns in the input photographs are subsequently detected using a classifier [5].

II. LITERATURE REVIEW

According to H. L. Choi and S. H. Kim et al. concentrate on utilizing Convolutional Neural Networks (CNNs) for detection of multiple targets. Convolutional Neural Networks (CNNs) is a deep learning model known for their effectiveness in tasks connected to images. In this context, the authors likely proposed a method that involves the use of CNNs to automatically detect and recognize multiple targets from aerial surveillance data [6].

According to C. L. Giles, S. Lawrence, et al, resents an early application of Convolutional Neural Networks (CNNs) for face recognition. In this pioneering work, the authors introduce a novel approach leveraging CNNs to address the complexities of face recognition tasks. The paper likely discusses the architecture of the CNN used, the methodology for training, and the application of this approach to a specific dataset for evaluation. Given the publication's timing, this work serves as a foundational contribution to the integration of neural networks, specifically CNNs, in the realm of facial recognition, paving the way for subsequent advancements in the field. The research is likely to have laid the groundwork for exploring the prospective of deep learning techniques in improving the accuracy and efficiency of face recognition systems [7].

Deep learning enables the training of extensive datasets by utilizing a hierarchical network model and employs a layer by layer feature extraction approach to acquire high-level image attributes. The deep learning network model is capable of extracting both the basic features and the deep features of an image by utilizing numerous hidden layers. Deep learning methods yield features that are not only accurate but also advantageous for picture categorization, in contrast to conventional machine learning techniques. The deep learning model primarily determines the method of feature learning and combination throughout the process of picture identification and classification [8]. At present, the most popular models in deep learning are convolution neural networks, sparse models, and restricted Boltzmann machines. Although both models have different feature extraction methods, they are extremely comparable when it comes to picture identification and classification.

An impactful strategy involves the utilization of deep learning models focused on feature representation, showcasing remarkable proficiency in the recognition and categorization of diverse image types. Among these models, deep convolutional neural networks (DCNs) have emerged as particularly notable. DCNs excel in extracting intricate visual features, rendering them highly effective when dealing with extensive datasets. This effectiveness is underscored by their success in large-scale datasets like ImageNet, where these models exhibit accurate classification of substantial volumes of visual data. The application of DCNs underscores the potency of deep learning techniques in augmenting the capacity of image classification models to handle complex visual information. Furthermore, the deep learning model demonstrates adeptness in autonomously learning and describing the inherent characteristics of images. It excels in acquiring hierarchical features even in the absence of human supervision, highlighting its autonomous learning capabilities. The model's recovered characteristics are highly expressive and not only make picture classification more efficient.. Extensive study has shown that deep learning models have gradually replaced traditional methods of machine learning and artificial feature extraction in image classification as research into related topics continues to grow in depth.[10].

III. PROPOSED METHODOLOGY

A. *Enhanced CNN*

Given the high real-time demands of systems utilizing image classification algorithms, it is imperative for these algorithms to prioritize real-time performance. Image classification using complicated neural network models requires a significant amount of time. Thus, this research streamlines the AlexNet model and utilizes it as the foundational model for picture categorization. [11].

One possible strategy is to use an automatic encoder with noise reduction capabilities in conjunction with a convolutional neural network. Improving picture classification and reducing the risk of over-fitting are both possible with data augmentation. When working with small data sets, this classification technique outperforms competing algorithms in terms of generalization performance. The method also includes an automatic encoder for noise reduction, which ensures the model can generalize well by reducing the impact of data noise on its performance. It is possible that the training time for the model will rise because the improved algorithm is built on top of the AlexNet network [12]. The image classification technique with optimization is shown in Figure 1.

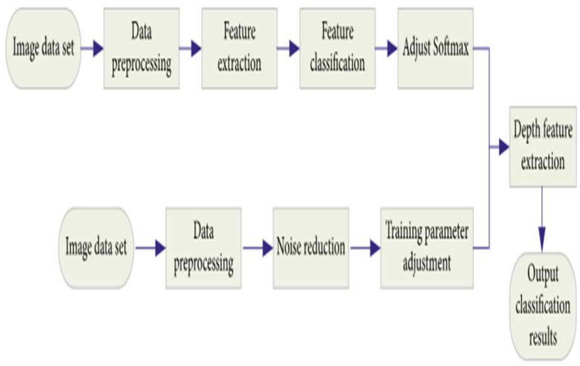

Fig 1. Image Classification Procedure

This research utilizes a normalized encoder network to address the issue of automatic noise reduction in intricate image structures. Integrating the noise reduction auto-encoder with the sparse auto-encoder enhances the convolutional neural network model. The sparse auto-encoder is utilized to standardize the initial image data. Finally, utilize the enhanced convolutional neural network model to extract feature data from the images, followed by the Softmax classifier to classify these features [13].

Prior to employing the enhanced convolutional neural network (CNN) model for image categorization, it is imperative to preprocess the images by reducing noise and converting them to grayscale. Additionally, the dataset should be systematically divided into predetermined training and test sets. Through an initial phase of unsupervised learning, the training set serves as input for the model. The automatic noise reduction encoder's hidden layer is then engaged in encoding and decoding the input data, followed by the utilization of the automated sparsity normalization encoder in the subsequent layer to process the data. Layer-wise training ensues, involving the hidden layer of a sparse auto-encoder. Ultimately, the training results from the sparse auto-encoder are fed into a Softmax classifier. Enhancing the deep learning models performance and accuracy in photo classification involves utilizing the gradient descent approach to train the parameters of the classifier model. The evaluation of the image classification technique is carried out using the picture test set, and the model's classification results are scrutinized for overall performance assessment

B. AlexNet

The inception of AlexNet points a crucial and groundbreaking development in the realm of deep learning, particularly designed for image categorization tasks. Unveiled by Ilya Sutskever, Alex Krizhevsky and Geoffrey Hinton in the year 2012. AlexNet introduced a notable architecture comprising eight layers. This architecture encompasses five convolutional layers, max-pooling, and three fully connected layers. Noteworthy features such as dropout regularization, Rectified Linear Unit (ReLU) activations, and Local Response Normalization (LRN) were incorporated, representing cutting-edge techniques. The integration of these techniques served dual purposes: enhancing training efficiency and mitigating the risk of over-fitting. The introduction of AlexNet was a significant milestone in advancing the capabilities of deep convolutional neural networks, particularly in the domain of image categorization.. One of the reasons it was successful in the 2012 ImageNet Large Scale Visual Recognition Challenge (ILSVRC) was because it was the first system of its type to employ GPUs to accelerate training. In addition to its obvious success, AlexNet is noteworthy for shaping the designs that followed it and, by extension, the current ecosystem of artificial intelligence applications. This, in turn, led to the extensive use of deep learning [15].

C. VggNet:

VGGNet, developed by the Visual Geometry Group at the University of Oxford, is recognized for its straightforward and consistent design. It was introduced by Simonyan and Zisserman in 2014. VGGNet's key characteristic is the use of very small (3x3) convolutional filters throughout the network, stacked together to form deeper architectures. The most common variants are VGG16 and VGG19, which differ in the number of layers. Despite its straightforward structure, VGGNet achieved competitive performance in image classification tasks and became a reference architecture for deeper networks [16].

D. LeNet:

LeNet, developed by Yann Lecun and his collaborators in the 1990s, is one of the earliest CNN architectures. It was designed for handwritten digit recognition and was instrumental in demonstrating the effectiveness of deep learning for image classification. LeNet comprises convolutional layers followed by average pooling and fully connected layers. It introduced key concepts such as convolutional layers with trainable parameters and shared weights, as well as sub sampling layers. Even if it doesn't measure up to modern designs, LeNet was essential in establishing deep learning and paving the way for advancements in CNN design. [17].

Before applying the enhanced convolutional neural network model for classification of images, it is crucial to preprocess the image by reducing noise and making it grayscale. It is also recommended to use the dataset to select a fixed amount of training and test sets. Following the preparation of the training data, the model will be subjected to unsupervised learning processes. Furthermore, the input entity is encoded and decoded utilizing the automated noise reduction encoder's hidden layer. The next layer's automated sparsity normalization encoder receives the procedure results. The image data is incrementally trained as layer by layer by utilizing hidden layer of a sparse auto-encoder. The Softmax classifier is employed to produce output based on the training outcomes of the sparse auto-encoder. One can enhance the deep learning model's performance in image classification by utilizing gradient descent method to optimize the training of the classifier model's parameters, resulting in more precise classifications. The model is evaluated using the image test set data and the model's classification results are used to assess the success of the image classification technique [20].

When it comes to image classification, traditional neural networks can only use a subset of available characteristics; this is something that the improved convolutional neural network model can overcome. Over-fitting can be avoided in data processing by using the normalizing sparse automated encoder technique. Another option is to use the sparse automated encoder's hidden layer for layer-by-layer data training; this will allow for the acquisition of representational and abstract features [18].

IV. RESULTS AND DISCUSSIONS

A. *Datasets*

Using the Flower dataset made available by Oxford University [19], we ran an experiment to assess the performance of the model for image classification presented in this study. Several flower species are shown in these photographs, along with their relative sizes, shapes, and lighting differences; certain floral categories show notable changes. There are a total of 1,360 photographs in the collection, spread among 17 different floral datasets. Each dataset contains 80 images. For this experiment, we randomly divide the dataset into three parts: one part for training the model, one part for validation, and one part for testing its performance.

The experiment utilized classification accuracy as a metric to measure the performance of the model. Evaluation of classification accuracy involves considering overall accuracy of the model and accuracy unique to each category. Overall accuracy of the model is based on ratio of correctly classified samples to the total number of samples, whereas accuracy of classification is measured as the ratio of correctly classified samples to the total number of tests samples. For the objective of conducting a comparison three widely used models, namely LeNet, AlexNet, and VggNet, have been chosen. Figure 2 illustrates the correlation between the classification accuracy and the number of iterations of the model while employing various network models, as indicated by the experimental comparison results. After performing 40 iterations, the LeNet model achieved an accuracy of 34%, the VggNet model achieved an accuracy of 39%, and the AlexNet model achieved an accuracy of 45%. The LeNet model achieves an accuracy of 45% after 80 iterations, whereas the VggNet model achieves an accuracy of 51% and the AlexNet model achieves an accuracy of 59%. Furthermore, upon convergence, the ECNN model achieves a remarkable accuracy rate of 96%, while the LeNet model reaches a maximum accuracy rate of 81%, and the VggNet model attains a peak accuracy rate of 85%.

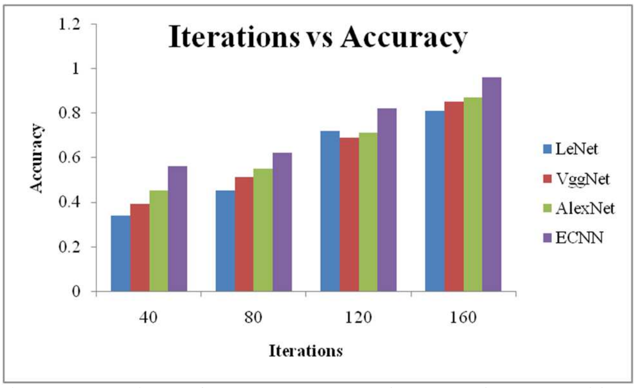

Fig 2. Comparison of LeNet, VggNet, AlexNet and ECNN with different iterations

To measure the comparison among loss value function with number of iterations both the training and testing datasets are considered.

To analyze the relationship between the loss value function of the optimized model and the number of iterations, the models are assessed using both the training set and the test set. This comparison is illustrated in Figure 3. The non-optimized model exhibits an increasing trend in the loss value function as the number of iterations increases, suggesting that it is prone to over-fitting. Conversely, the optimized model demonstrates a decreasing trend in the loss value function as the number of iterations increases. Evidently, model optimization can effectively decrease the expense associated with parameter training.

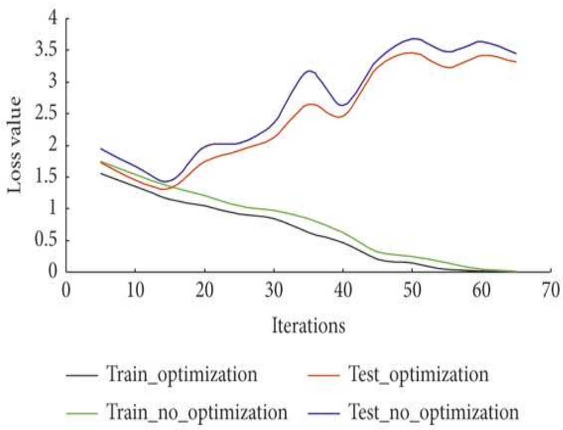

Fig 3. Relationship between loss value with number of iterations

The study's proposed model has superior classification accuracy compared to previous models, as indicated by experimental results analyzing the correlation between performance of traditional models in picture classification and the number of repetitions. Image classification accuracy can be greatly enhanced with optimization, as demonstrated by comparing the deep learning model's classification accuracy on the training set and the test set before and after optimization.

V. CONCLUSION AND FUTURE SCOPE

This paper introduces a deep learning model for image classification that aims to overcome the challenges of long processing time and low accuracy often encountered in traditional image classification methods. The proposed model relies on machine learning. This study explains the basic principles of neural networks and explores various types of convolutional neural networks, focusing on their use in image classification. A new deep learning model for image classification was developed by implementing noise reduction and parameter control techniques in the feature extraction stage. This research aimed to improve the effectiveness and precision of a new deep learning model based on an enhanced convolutional neural network. The main goal was to enhance the classification accuracy of the model. Experimental studies were performed to evaluate the precision of different widely-used network models for image categorization. The findings clearly showed that the proposed model outperformed its competitors in terms of categorization accuracy. Comparative evaluations were conducted to assess the classification accuracy of the deep learning model before and after optimization. The results showed a notable enhancement in image categorization accuracy using the optimized model. In the future, research should focus on improving current methods for dynamic target classification in intricate contexts, using the knowledge obtained from this study as a foundation.

REFERENCES

[1] P. W. Song, H. Y. Si, H. Zhou, R. Yuan, E. Q. Chen, and Z. D. Zhang, "Feature extraction and target recognition of moving image sequences," IEEE Access, vol. 8, pp. 147148–147161, 2020.

[2] Venkata Nagaraju Thatha; A. Sudhir Babu; D. Haritha, Privacy-preserving smart contracts for fuzzy WordNet-based document representation and clustering using regularised K-means method. International Journal of Ad Hoc and Ubiquitous Computing, 2022 Vol.40 No.1/2/3, pp.2 – 9. DOI: 10.1504/IJAHUC.2022.123521.

[3] W. Y. Zhang, X. H. Fu, and W. Li, "The intelligent vehicle target recognition algorithm based on target infrared features combined with lidar," Computer Communications, vol. 155, pp. 158–165, 2020.

[4] M. Li, H. P. Bi, Z. C. Liu et al., "Research on target recognition system for camouflage target based on dual modulation," Spectroscopy and Spectral Analysis, vol. 37, no. 4, pp. 1174–1178, 2017.

[5] Fei-Fei Li, Justin Johnson and Serena Yueng, "Lecture 9: CNN Architectures" May 2017.

[6] Khan, S. M. H., Hussain, A., &Alshaikhli, I. F. T. (2012, November) "Comparative study on content-based image retrieval (CBIR)." in Advanced Computer Science Applications and Technologies (ACSAT), 2012 International Conference. .

[7] B. VeeraSekharReddy, K. S. Rao and N. Koppula, "Named Entity Recognition using CRF with Active Learning Algorithm in English Texts," 2022 6th International Conference on Electronics, Communication and Aerospace Technology, Coimbatore, India, 2022, pp. 1041-1044.

[8] W. L. Xue and T. Jiang, "An adaptive algorithm for target recognition using Gaussian mixture models," Measurement, vol. 124, pp. 233–240, 2018.

[9] F. Liu, T. S. Shen, S. J. Guo, and J. Zhang, "Multi-spectral ship target recognition based on feature level fusion," Spectroscopy and Spectral Analysis, vol. 37, no. 6, pp. 1934–1940, 2017..

[10] S. Razakarivony and F. Jurie, "Vehicle detection in aerial imagery: a small target detection benchmark," Journal of Visual Communication and Image Representation, vol. 34, pp. 187–203, 2016.

[11] W. Z. Wu, J. W. Zou, J. Chen, S. Y. Xu, and Z. P. Chen, "False-target recognition against interrupted-sampling repeater jamming based on integration decomposition," IEEE Transactions on Aerospace and Electronic Systems, vol. 57, no. 5, pp. 2979–2991, 2021.

[12] T. Yulin, S. H. Jin, G. Bian, and Y. H. Zhang, "Shipwreck target recognition in side-scan sonar images by improved YOLOv3 model based on transfer learning," IEEE Access, Z. Geng, H. Deng, and B. Himed, "Ground moving target detection using beam-Doppler image feature recognition," IEEE Transactions on Aerospace and Electronic Systems, vol. 54, no. 5, pp. 2329–2341, 2018.

[13] Bhumireddypalli, VSR, Koppula, SR, Koppula, N. "Enhanced conditional random field-long short-term memory for name entity recognition in English texts", in Concurrency Computation: Practice and Experience. 2023; 35(9):e7640. doi: 10.1002/cpe.7640

[14] V. N. Thatha, B. VeerasekharReddy, G. VenuGopal, K. Ashok, S. Maddu and S. V, "An Enhanced Support Vector Machine Based Pattern Classification Method for Text Classification in English Texts," 2023 7th International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS), Bangalore, India, 2023, pp. 1-6.

[15] S, Phalguna &Thatha, Venkata&Mamidisetti, Gowtham & Mantena, Srihari & Chintamaneni, Phanikanth & Vatambeti, Ramesh. (2023). Hybrid deep learning model with enhanced sunflower optimization for flood and earthquake detection. Heliyon. 9. e21172.

[16] VeeraSekharReddy, B., Rao, K.S. &Koppula, N. "An Attention Based Bi-LSTM DenseNet Model for Named Entity Recognition in English Texts", in Wireless Personal Communications 130, 1435–1448 (2023).

[17] Z. -H. Zhang, Z. Yang, Y. Sun, Y. -F. Wu and Y. -D. Xing, "Lenet-5 Convolution Neural Network with Mish Activation Function and Fixed Memory Step Gradient Descent Method," 2019 16th International Computer Conference on Wavelet Active Media Technology and Information Processing, Chengdu, China, 2019, pp. 196-199.

[18] B. Veerasekhar Reddy, K. Srinivas Rao, K. Neeraja; "Named entity recognition on different languages: A survey". AIP Conference Proceedings 22 May 2023; 2492 (1): 030043.

[19] M. E. Nilsback and A. Zisserman, "Automated flower classification over a large number of classes," in Proceedings of the Sixth Indian Conference on Computer Vision, Graphics & Image Processing, pp. 722–729, Bhubaneswar, India, December 2008.

[20] C. Y. Zhang, B. L. Guo, N. N. Liao et al., "A public dataset for space common target detection," KSII TRANSACTIONS ON INTERNET AND INFORMATION SYSTEMS, vol. 16, no. 2, pp. 365–380, 2022.