Deep Learning-Based Image Classification

Dharshini S[1], Dr C Thiyagarajan[2]
[1]Dept of MCA
[2]Associate Professor, Dept of MCA
[1, 2] PSG College of Arts & Science, Coimbatore, India

Abstract- *Image classification serves the purpose of narrowing the divide between computer vision and human perception, striving to enable machines to recognize images in a manner akin to human understanding. Although traditional systems have found widespread application in practical scenarios, they suffer from several issues, including low classification accuracy, limited adaptability, and suboptimal results. Consequently, deep learning has emerged as a solution for image classification, offering the potential to mitigate these problems and enhance accuracy.*

This paper delves into the foundational concepts related to image classification using deep learning, explores various algorithms employed for this task, discusses the advantages of deep learning in image classification, and investigates the diverse applications where deep learning-based image classification finds utility.

In summary, this paper provides an extensive examination of image classification with deep learning, covering its significance, shortcomings of traditional approaches, the pivotal role of deep learning, and its real-world applications. This research area has profound implications across industries and has the potential to revolutionize computer vision capabilities.

I. INTRODUCTION

According to IDC, it was projected that by 2020, the global data volume would reach an astonishing 45ZB, with approximately 70% of this data being composed of images or videos. To harness valuable insights from this wealth of visual data, the field of computer vision has emerged as a pivotal player.

In the current landscape, computer vision has undergone rapid advancements, particularly in the realm of image classification. The roots of image classification can be traced back to the late 1950s, and since then, its applications have extended across a wide array of domains, encompassing engineering, medicine, human-vehicle tracking, fingerprint analysis, disaster management, and security.

At its core, classification involves the systematic categorization of data into distinct groups and categories, contingent upon discernible features. The emergence of image classification was driven by the aspiration to bridge the divide between computer vision and human vision by instructing computers through data. In the context of image classification, an image is categorized by distinguishing it into predefined categories based on its visual content.

Numerous methodologies for image classification have been proposed, and they can be broadly classified into four categories:

- Statistics-based Image Classification: This methodology leans on statistical models, such as the Markov model and Bayesian model, with the objective of minimizing classification errors.
- Texture, Local Features, and Traditional Colors-based Image Classification: This approach entails the scrutiny of image textures, local features, and traditional color patterns for the purpose of classification.
- Deep Learning-based Image Classification: Deep learning, a subfield of machine learning, has garnered significant attention. It empowers computers to autonomously learn and classify images, sounds, and text. In the realm of deep learning, a computer is trained using extensive image datasets, transforming pixel values into internal representations, subsequently enabling classifiers to detect patterns within input images.

The conventional approach to image classification hinges on machine learning, which typically encompasses a feature extraction module that identifies crucial features, such as edges and textures. This is followed by a classification module that assigns labels based on these extracted features. Nevertheless, machine learning is not without its limitations; it can only extract features resembling those found in the training data and is unable to adapt to novel features.

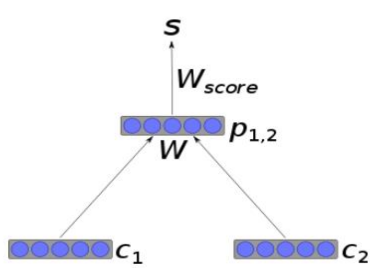
To surmount these constraints, deep learning emerged as a subfield within machine learning. Deep learning algorithms possess the capability to self-learn through their proprietary computational methods. In the realm of deep learning, a computer is educated to classify images, sounds, and text by undergoing training on extensive image datasets. Subsequently, it transforms pixel values into internal representations, and classifiers then discern patterns within input images.

One of the most frequently employed deep learning methods for image classification is the Convolutional Neural Network (CNN). CNNs offer several advantages, including parameter sharing, equivalent representation, and sparse interactions, rendering them highly effective for image classification tasks. The various networks of deep learning methods are as follows:-

The **Convolutional Neural Network (CNN)** stands out as a widely employed deep learning technique for image classification. It possesses a distinct advantage – the ability to directly glean insights from image data, rendering manual feature extraction unnecessary. CNNs offer three key merits: parameter sharing, equivalence in representation, and sparse interactions.

In the sphere of deep learning, the **Recurrent Neural Network (RNN)** holds its own, particularly in fields like Natural Language Processing (NLP) and speech processing. RNNs leverage sequential data and can be envisioned as a form of short-term memory, comprising concealed layers, input layers (x), and output layers (y).

The **Recursive Neural Network (RvNN**) is equipped to make predictions in a hierarchical structure and perform object classification using compositional vectors. This is achieved by iteratively applying the same set of weight parameters to structured inputs, resulting in structured predictions that accommodate variable-sized input structures.



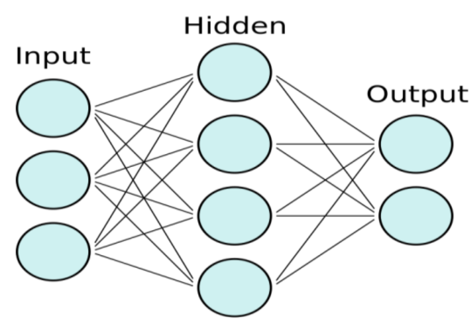
A SIMPLE RvNN ARCHITECTURE

A CNN is constructed with several layers:

- Convolutional layer: This layer conveys information to the subsequent layer, facilitating the extraction of distinctive features.
- Pooling layer: It consolidates clusters of neurons into single neurons in the next layer, thereby reducing dimensionality.
- Fully connected layer: This layer interconnects each neuron in one layer with every neuron in another, enabling comprehensive learning and facilitating classification.

In the CNN paradigm, the acquisition of feature detection capabilities unfolds through the utilization of numerous hidden layers, typically numbering in the tens or even hundreds. Each of these layers augments the complexity of acquired features. This hierarchical structure empowers CNNs to discern intricate patterns and representations within the data, thus enabling proficient image classification.

The process typically begins with an input image, where various filters are applied to create a feature map. Following this, the Rectified Linear Unit (ReLU) function is applied to enhance non-linearity. Subsequently, a pooling layer is applied to each feature map, and the pooled images are flattened into a single, elongated vector. This vector is then fed into a fully connected artificial neural network. The network is trained through forward and backward propagation iterations, repeating until it possesses well-defined neural network architecture with trained weights and feature detectors.



When considering the application of Convolutional Neural Networks (CNNs) for image classification in greater detail, the primary objective is to accept an input image and assign it to its respective class. While humans naturally classify images, computers perceive images differently.

Fig1.1:What we see

151	121	1	93	165	204	14	214	28	235
62	67	17	234	27	1	221	37	189	141
20	168	155	113	178	228	25	130	139	221
236	136	158	230	10	5	165	17	30	155
174	148	93	70	95	106	151	10	160	214
103	126	58	16	138	136	98	202	42	233
235	103	52	37	94	104	173	86	223	113
212	15	179	139	48	232	194	46	174	37
119	81	241	172	95	170	29	210	22	194
129	19	33	253	229	5	152	233	52	44
88	200	194	185	140	200	223	190	164	102
113	16	220	215	143	104	247	29	97	203
9	210	102	246	75	9	158	104	184	129
124	52	76	148	249	107	65	216	187	181
6	251	52	208	46	65	185	38	77	240
150	194	28	206	148	197	208	28	74	93
33	183	248	153	168	205	146	100	254	218
130	53	128	212	61	226	201	110	140	183
165	246	22	102	151	213	40	138	8	93
152	251	101	230	23	162	70	238	75	24
187	105	152	83	167	98	125	180	136	121
139	197	55	209	28	124	208	208	104	40
123	19	144	223	62	253	202	108	47	242
220	144	31	16	136	123	227	62	183	163

Fig1.2:What a computer see

To a computer, an image is an array of pixels, for instance, in a 200x200 image, the array size would be 200x200x3, where 200 represents the width, 200 is the height, and 3 accounts for RGB values.

To address this disparity, computers look for basic characteristics at the lowest level, such as edges or curvatures. The process commences with the convolution layer, where the image is initially inputted. Subsequently, software selects a smaller matrix known as a filter, which performs convolution by multiplying its values with pixel values and summing up the results. This output from one convolution layer becomes the input for the next, repeating through each layer.

Following the convolution layer, a nonlinear layer is introduced. This layer incorporates an activation function, introducing nonlinearity, which is crucial for the network's capacity to model response variables effectively.

The pooling layer follows the nonlinear layer and operates on the image's width and height, downsampling them to reduce image volume.After these layers, a fully connected layer is appended, resulting in an N-dimensional vector, with N representing the number of classes from which the model selects the appropriate class.

The construction of the model begins with defining the object, for instance, model = Sequential(), and then adding layers with their respective types using model.add(type_layer()). After the addition of sufficient layers, the model is compiled, specifying a loss function and optimizer algorithm.The next step is model testing, employing a dataset that has not been used previously to assess accuracy. Once the training is complete and the model yields accurate results, it can be saved using model.save("name_of_file").

Ultimately, the model can be deployed in the real world. This phase, known as model evaluation, involves utilizing the model to assess new data.

II. RESULT

Upon testing the model, it achieved an impressive accuracy of 96%. However, a noteworthy issue was the extended duration required for the training process, with approximately 50 minutes needed to complete 60 epochs.

Let's consider the creation of a plot to visually represent our achieved accuracy of 96% in image classification, which is indeed a commendable result.

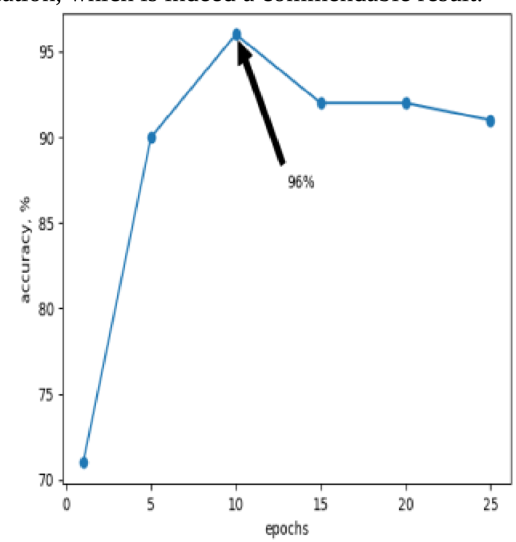


Fig2.1:Accuracy of the model(%)

III. LITERATURE REVIEW

In a research conducted by Mohd Azlan Abu and Nurul Hazirah Indra in 2019 [1], they utilized the TensorFlow framework for image classification through deep neural networks. Python was selected as the programming language due to its seamless integration with TensorFlow. Their dataset consisted of thousands of flower images, encompassing five distinct flower types, with each category comprising hundreds of images. The deep neural network (DNN) underwent extensive training on this dataset until it exhibited the capability to recognize all the images within it. Subsequently, successful image classification was achieved, resulting in an impressive accuracy rate of approximately 90%.

In their 2018 study, Karan Chauhan and Shrwan Ram [2] explored image classification using a deep learning approach. They employed a convolutional neural network (CNN) based on Keras and TensorFlow in Python. Their research compared four distinct CNN structures on a CPU system, each with different combinations of classifiers (softmax and sigmoid) and activation functions (ReLU and Tanh). They used a dataset containing 10,000 dog and cat images, with 8,000 for training and 2,000 for testing. After defining parameters and creating a CNN with two convolutional layers, they trained the model. The outcome was a highly accurate image classification system.

In their 2017 study, Dr. Vinayak Bharadi and Misbah Naimuddin Panchbhai [3] introduced a system called "image classification using deep learning." This system employs neural networks for image classification and can assess classification accuracy on both GPU and CPU platforms using Python. The process involves capturing and normalizing images, reducing dimensionality with techniques like Block Transition Coding and Histogram analysis. Performance is evaluated based on metrics like accuracy and execution time, allowing for a comparison between CPU and GPU efficiency.

In their 2018 research [4], M. Manoj Krishna and M. Neelima utilized the AlexNet architecture in combination with a convolutional neural network (CNN) for the purpose of image classification. They selected four test images from a database for their classification experiments, and these images were further divided into various segments. Their methodology involved training the neural network by presenting it with an input image and providing it with the expected output. Notably, the results of their experiments demonstrated the algorithm's capability to accurately classify images, even when dealing with partial test images. This underscores the effectiveness of deep learning algorithms in the domain of image classification.

In their 2020 publication [5], W. A. Ezat, M M Desouky, and N A Ismail utilized pre-trained CNN models sourced from ImageNet to classify images within the PASCAL VOC dataset. They adopted a transfer learning approach to enhance the deep learning CNN model's performance, achieving effective classification even with limited computational resources and minimal computation time. Their research outcomes were juxtaposed with results obtained through alternative methods, including SVM, region ranking SVM, and super vector coding of local image descriptors. The paper provides two significant contributions. Firstly, it presents a comprehensive comparison between deep learning algorithms for multi-class image classification and other classification techniques, assessing them based on their performance. Secondly, it underscores the ease of implementing deep learning algorithms in real-world applications.

In their 2019 paper [6] titled "Image Classification Algorithm based on Deep Learning – Kernel Function," Jun-e-Liu and Feng-Ping An introduced an innovative approach to image classification. Their method involves a clear separation of image feature extraction and classification into two distinct steps. A notable aspect of their work is the integration of sparse representation into the architecture of deep learning networks. This incorporation of sparse representation addresses the challenge of approximating complex functions and results in the development of a deep learning model with adaptive approximation capabilities. In summary, their approach enhances the adaptability and efficacy of deep learning models for image classification by incorporating sparse representation and dividing the process into separate feature extraction and classification phases.

In their 2018 project [7], Muthukrishnan Ramprasath and M. Vijay Anand introduced a system titled "Image Classification using Convolutional Neural Network." This system harnesses deep learning algorithms to attain precise results in computer vision tasks. Specifically, it employs Convolutional Neural Networks (CNNs) to autonomously classify images. To evaluate its performance, they used MNIST digits as a standard benchmark for grayscale image classification. By training their data using a CNN network, they achieved an impressive accuracy rate of 98%. The training process encompassed four distinct phases, including data pre-processing, feature extraction, data normalization, and image classification using Support Vector Machines (SVM). In summary, their system highlighted the potential of CNNs and SVMs for accurate image classification, particularly in the context of grayscale images, achieving a remarkably high level of accuracy.

In their 2019 research project [8], Mingyuan Xin and Yong Wang presented a study titled "Image Classification Model based on Deep Convolution Neural Network." Their investigation delved into the intricacies of the backpropagation algorithm, with a specific focus on training deep neural networks to minimize classification errors to the fullest extent. They explored three distinct techniques: the utilization of a pre-trained CNN model, the implementation of a hybrid approach, and the application of fine-tuning methods. The combination of pre-trained CNN and fine-tuning was seamlessly incorporated into the network, while the hybrid method involved feature extraction based on image patches. The primary benefits of these approaches manifested in improved speed and performance. These methods were demonstrated to enhance the efficiency and accuracy of image classification tasks, underlining the advantages of their chosen methodologies.

Table 1- Advantages of the corresponding conference

Reference	Advantages
Mohd Azlan Abu, Nurul Hazirah Indra(2019) [1]	1- 90% of accuracy was found .
Karan Chauhan, Shrwan Ram(2018) [2]	1- The results displayed a commendable level of accuracy. 2- The model exhibited rapid training, consuming minimal time for completion.
Dr. Vinayak Bharadi, Misbah Naimuddin Panchbhai(2017) [3]	1- The system possesses the capacity to evaluate the precision of image classification on both CPU and GPU processing units.
M Manoj Krishna, M Neelima(2018) [4]	1- The images were correctly classified, even when considering a fraction of the test images.
W. A. Ezat, M M Desouky and N A Ismail(2020) [5]	1- It highlights the straightforwardness of implementing deep learning algorithms. 2- Performance comparisons were provided as part of the analysis.
Jun-e-Liu, Feng-Ping An(2019) [6]	1- The problem of approximating a complex function was successfully resolved.
Muthukrishnan Ramprasath, M. Vijay Anand(2018) [7]	1- Impressive accuracy of nearly 98%.
Mingyuan Xin and Yong Wang(2019) [8]	1- Better speed and performance

IV. CONCLUSION

In conclusion, our exploration of various image classification techniques has revealed that each approach possesses its distinct advantages. Some prioritize accuracy, others prioritize performance, and some are cost-effective. The field of image classification has made substantial strides in recent times, but there remains ample room for further improvement. These advancements encompass diverse areas, including algorithm sophistication, dataset diversity, and specialized hardware utilization. As a result, image classification is a dynamic field with promising prospects for continued evolution and refinement in the future.

REFERENCES

[1] Mohd Azlan Abu, Nurul Hazirah Indra(2019) [1]- Image classification using deep neural network.

[2] Karan Chauhan, Shrwan Ram(2018) [2]- Image classification using deep learning based on keras and tensorflow.

[3] Dr. Vinayak Bharadi, Misbah Naimuddin Panchbhai(2017) [3]- Image classification using deep learning.

[4] M Manoj Krishna, M Neelima(2018) [4]- Image classification using deep neural network.

[5] W. A. Ezat, M M Desouky and N A Ismail(2020) [5]- Image classification using convolutional neural network.

[6] Jun-e-Liu, Feng-Ping An(2019) [6] Image Classification Algorithm based on Deep Learning – Kernel Function.

[7] Muthukrishnan Ramprasath, M. Vijay Anand(2018)[7] Image Classification using Convolutional Neural Network.

[8] Mingyuan Xin and Yong Wang(2019) [8] Image Classification Model based on Deep Convolution Neural Network.

[9] https://in.mathworks.com/matlabcent ral/fileexchange/59133-neuralnetwork-toolbox-tm--model-foralexnet-network.

[10] Diligenti, M.; Gori, M.; and Sacc`a, C. 2015. Semanticbased regularization for learning and inference.