International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 12, No.1 (Dec-2022)

https://dx.doi.org/10.12785/ijcds/1201109

Recent Researches on Image Classification Using Deep Learning Approach

Nayan Kumar Sarkar[1], Moirangthem Marjit Singh[1] and Utpal Nandi[2]

[1]*Department of Computer Science and Engineering, North Eastern Regional Institute of Science and Technology, Arunachal Pradesh, India*
[2]*Department of Computer Science, Vidyasagar University, Midnapore, West Bengal, India*

Received 10 Oct. 2021, Revised 14 Sep. 2022, Accepted 1 Dec. 2022, Published 8 Dec. 2022

Abstract: Image classification is an essential and widely used area where deep learning is applied. The deep learning approach has been extensively applied in the area of image classification and provides very good classification accuracy. Some of the deep learning approaches can classify images better than a human. Image classification has tremendous applications in practice. This paper presents a survey on some of the deep learning approach-based image classification methods which have been extensively used in various applications of image classifications. The deep learning approaches which have been considered in our study and are used for developing a variety of image classification methods are Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Long Short Term Memory (LSTM), Generative Adversarial Networks (GAN), Restricted Boltzmann Machine (RBM) and Deep Belief Network (DBN). The paper also discusses comparative studies on some of the image classification techniques which have been used in different areas of image classification problems.

Keywords: Image Classification, CNN, RNN, LSTM, GAN, RBM, DBN

1. Introduction

Deep Learning (DL) is a popular sub-field of machine learning (ML). It mimics the human brain, that is, how a human brain works. DL comprises one input layer, more than one hidden layer, and an output layer. It is a very effective, efficient, and fastest-growing field of ML. It has been applied to numerous fields, such as image classification, natural language processing, computer security, healthcare, self-driving car, robotics, etc. DL approaches provide outstanding results in various fields of image classification. For example, R. A. Hazarika et al., [1] in their study found that DL based classification approaches provide very good results in classifying Alzheimer's disease (AD).

Image classification is a widely used area where DL approaches are applied. It is also considered a major part of developing scientific equipment. Generally, good performance or classification accuracy is considered as one of the major parameters to develop a piece of equipment which helps to make decisions. Therefore, an image classification method with very good classification accuracy and less complexity takes a lead role in developing scientific equipment such as automated disease detectors, like breast cancer detector, COVID-19 detector, tumor detector, autonomous road crack detection system, etc. Hence, researchers have been working on developing image classification methods for improving the performance as well as development of various methods relying on different DL approaches.

The motivation of this paper is due to the advances in technology and computing paradigms especially in deep learning during recent times. Also, it is necessary to look towards techniques for achieving better performance in applications involving image classification. The paper focuses on the survey of DL approach based image classification techniques that have been developed over time. We highlight the different image classification techniques that were developed using either a single or multiple DL approach. Different image classification applications covered in the paper are breast cancer detection, brain tumor detection, COVID-19 detection, road crack detection, hyperspectral image (HSI) classification, eye disease classification, etc. It is found that for a specific classification problem many researchers have developed different methods based on different DL approach. For example, researchers have developed methods for analyzing the classification performance of hyperspectral images, breast cancer detection, tumor detection, etc. Section 2-7 highlights the CNN, RNN, LSTM, GAN, RBM, and DBN -based image classification approaches respectively in the paper. Section 8 shows the performance comparison of various approaches used in various applications. Section 9 presents a brief summary

E-mail address: nayankrsarkar@gmail.com, marjitm@gmail.com, nandi.3utpal@gmail.com

of the methods surveyed. Conclusion and future work are given in section 10.

2. Convolutional Neural Network (CNN)

CNN is a biologically inspired deep learning model. Generally, the architecture of CNN consists of several convolutions and pooling layer that is followed by fully connected layers. The convolution layer serves as a feature extractor. It consists of different shaped filters that compute different feature maps. The output of the convolution layer can be expressed by the following equation:

$$O_j^l = f\left(\sum_{i=1}^{N} P_j^{l-1} * Q_i^{l-1} + b_j^l\right) \quad (1)$$

where, O_j^l is considered as j^{th} feature map of layer l, P_j^{l-1} is j^{th} filter of $l-1^{th}$ layer. Q_i^{l-1} is i^{th} feature map of layer $l-1$ and b_j^l is considered as bias of j^{th} feature map of layer l [2].

The pooling layer is used to reduce the resolution of the feature maps after the convolution layer. It is applied between two convolution layers. The feature maps of the final layer are passed to the fully connected layer for classification. From the extracted features, the fully connected layer computes the classification score of each class and the highest score is considered as the resulted class. Depending upon its architecture, there can be multiples convolution layers followed by a pooling layer and finally a fully connected layer. There can be one or more fully connected layers. The general architecture of a CNN model is shown in Figure 1. The convolution and pooling layer is called as feature extraction layer and the fully connected layer is called as classification layer.

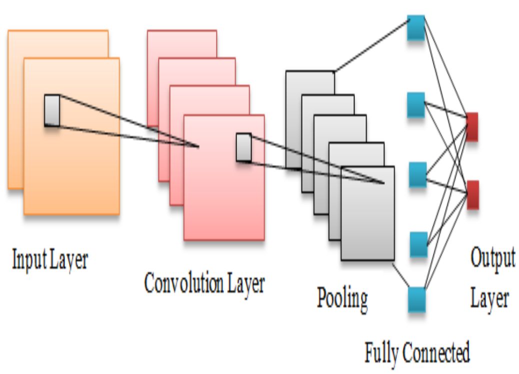

Figure 1. An overview of CNN architecture

The first model of CNN called ConvNet was developed by Y. LeCun et al., [3]. Later, various models like LeNet Y. LeCun et al., [4], AlexNet A. Krizhevsky et al., [5], VGG K. Simonyan et al., [6], GoogleNet C. Szegedy et al., [7], ResNet K. He et al., [8], and many more have been developed. The architecture of some developed models is very similar to each other but with a deeper layer. The popularity of the CNN model increased after the development of AlexNet architecture which showed the outstanding result in 2012-ILSVRC with a reduction of error rate from 25.8% to 16.4% in comparison to other conventional techniques [9]. CNNs have been successfully applied in image classification, speech recognition, NLP, handwriting recognition, face detection, etc. Among the various applications, CNNs are very famous for their application in image classification.

Numerous researches on image classification have been done based on CNN. X. Yang et al., [10] have proposed a deep CNN (DCNN) based method for automatic tumor tissue classification. The performance of their proposed model was superior in comparison to the support vector machine (SVM) and extreme learning model (EML). Their proposed model was general and applicable to other biological datasets. K. M. Hosny et al., [11] have proposed an automated skin lesion classification method. They applied the transfer learning paradigm and data augmentation techniques to AlexNet to overcome the requirement of huge labeled images. They used the ph2 dataset for their work and the performance was compared with other methods by considering the quantative measures, accuracy, sensitivity, specificity and precision. S. Deepak et al., [12] have suggested a method for classifying brain tumors. Their system adopted a transfer learning technique and used pre-trained GoogleNet for extracting brain MRI features automatically. Their method achieved 98% classification accuracy and also addressed the situation of calculating performance with less training data. H. Moujahid et al., [13] have proposed a CNN-based automated method for the detection of pneumonia from X-ray images. In their method, with CNN they have used the concept of transfer learning to detect pneumonia and also compared the result with other models for detecting the best model. P. Yamlome et al., [14] have developed a method for improving the result of CNN-based classifiers for the identification of breast cancers. To increase the result of the CNN classifier, they offered an effective training approach that comprises a combination of transfer learning with data augmentation and the entire image. Their method achieved significant improvement in comparison to other reported methods. S. Srinivasan et al., [15] proposed DCNN based image spam classification method. They considered two DCNN in their work. Each CNN contains 3 convolutional layers, ReLU activation function with a variety of filter sizes, and max-pooling size. In their work, the DCNN models are trained on three distinct datasets- image spam hunter dataset, improved dataset and dredze image spam dataset. Z. Chen et al., [16] proposed a method named Virtual Pooling (ViP) for improving the speed and consumption of energy of the CNN-based methods which are developed for the tasks of classification of images and detection of objects. They have shown the efficacy of ViP on four types of CNNs with three typical datasets on both mobile and desktop platforms and also for the task of image classification and object detection.

Many researchers have used CNN for hyperspectral remote sensing image classification. H. Chen et al., [17] and F. Zhang et al., [18] have proposed CNN-based hyperspectral image classification approaches. M. F. Aslan et al., [19] proposed mAlexNet and BiLSTM-based hybrid method for covid-19 detection. In recent works, V. K. Waghmare et al., [20] have implemented basic CNN and VGG-16 architecture for automatic brain tumor classification. Their model works in three different steps - (a) preprocessing step for reducing noise and artifacts (b) data augmentation step for increasing the size of the dataset and finally (c) classification step by using fine-tuned VGG-16 CNN. Their finding could be applied as an important part of a computer-aided diagnosis tool that could be used by clinicians for identifying brain tumors and their types. A. Anitha et al., [21] proposed a CNN-based automated technique for classifying eye diseases from the fundus image. Their approach provides outstanding performance and has been embedded in a hospital management system for performing automatic diagnosis on fundus images. J. Gross et al., [22] have proposed an automatic and very effective CNN-based image classification approach for identifying various types of Corneal Ulcers from fluorescein staining images. Their approach provided 92.73 percent balanced accuracy and also can be contributed to the IT-based healthcare application.

Md. A. Islam et al., [23] have proposed a CNN-based automated approach for paddy leaf disease detection. Their model could detect 4 types of diseases among 30 types of different paddy leaf diseases. They have analyzed the performances of four CNN architectures- Inception-Resnet-V2, VGG-19, Xception, and ResNet-101 and reported Inception-Resnet-V2 architecture provides the highest accuracy. The key objective of their research was the detection of diseases of paddy leaves more accurately and quickly in comparison to traditional time-consuming methods where the accuracy is also doubtful. A. Narin et al., [2] have developed an automated COVID-19 disease detection system based on five pre-trained CNN architectures and by considering chest X-ray images. The CNN models they have utilized in their research are ResNet50, ResNet101, ResNet152, InceptionV3 and Inception-ResNetV2. The X-ray images they have used for performance evaluation were obtained from normal, COVID-19 positive patients, bacterial and viral pneumonia patients. The result of their study shows that among the five models ResNet-50 provides the highest accuracy with the used datasets. It is believed that their model would be helpful for radiologists in making decisions in clinical practice. Based on CNN N. Kesav et al., [24] have suggested a simple model for brain tumar classification. R. Amalia et al., [25] have proposed a CNN + LSTM based approach for detecting and generating a description of diabetic retinopathy from fundus images. CNN was used for detecting lesions on retinal fundus images and LSTM was used for the generation of descriptions in sentential form based on the detected lesions. The proposed model achieved 90 percent accuracy and could be helpful for radiologists in analyzing diabetic retinopathy. K. E. Asnaoui [26] has developed an ensemble model consisting of three CNN architectures for the classification of pneumonia disease using X-ray images. The three CNN architectures he had used in his model were ResNet50, Inception-ResNet-V2, and MobileNet-V2. The result states that the developed ensemble model is superior in comparison to other ensemble models.

3. Recurrent Neural Network (RNN)

RNN is another deep learning model which remembers the current input it receives from the previous step and the remembered input allows it to predict the output. RNN is a feedforward neural network that has internal memory. RNN uses its internal memory for processing the sequence of inputs. In RNN, the output of the current layer becomes the input to the next layer. At a time state t, with the given previous layer's state h_{t-1}, the current and output layer's state C_t and O_t can be calculated respectively by the followings:

$$C_t = a_h \left(B_h I_t + C_h h_{t-1} + b_h\right) \quad (2)$$

$$O_t = a_y \left(B_y C_t + b_y\right) \quad (3)$$

Where, I_t is a feature vector, B_h and B_y are weight matrices of input-hidden and hidden-output layer. C_h is the recurrent weights matrix, b_h, b_y and a_h, a_y are the biases and activation functions respectively. In other deep learning models, the inputs are independent of each other but in RNN the input is related to each other. The problem faced by RNN is the Vanishing gradient problem. RNN is specially used in Natural Language Processing (NLP) and speech processing [27]. It has been used in image classification also.

Several image classification approaches based on the RNN model have been developed. L. Mou et al., [28] have developed a RNN based method for classifying hyperspectral images. They also introduced an activation function called PRetanh for preprocessing of hyperspectral data in RNN. Further, they have introduced a modified GRU with PRetanh for the effective analysis of hyperspectral data. Their method provided better hyperspectral image classification accuracy in comparison to CNN and SVM-RBF methods. L. Li et al., [29] have worked on an end-to-end RNN model for the weakly supervised multi-level image classification. The developed approach was comprised of three modules- recurrent highlight network (RHN), emission indicator, and gated recurrent relation extractor (GRRE). Each of these components performed its respective tasks. They evaluated the performance of their approach by using the NUS-wide, MS-COCO, and ESP Game dataset. The results of the evaluation of the model on these datasets stated that the proposed model overtook the performance of other state-of-the-art models and also performed better on categorizing small objects with a large number of labels. X. Zhang et al., [30] have worked on improving RNN based method to classify hyperspectral image. Proposed models are called LSS-RNN and NLSS-RNN. They used three different datasets such as Indian pines, Salinas valley, and the University of Pavia dataset for performance evaluation and the result proved that their methods were better than other state-of-the-art methods.

F. Lyu et al., [31] presented a method to classify multi-label images with visual attention and RNNs. The classification of an image with multiple-label is generally more challenging in comparison to classifying an image with a single label, because-firstly, the multi-label image consists of numerous objects and which may be in any place of the image, secondly, the significance of dissimilar areas of the image are dissimilar and lastly, multiple labels can have label dependency. Their proposed method solved all these three challenges. To establish the efficiency of their model, they had used two datasets MS-COCO, and pascal-VOC, and the result showed that their method performed better than other modern methods. R. Venkatesan et al., [32] proposed an RNN based approach for classifying hyper-spectral images. Their model analyzed pixels of hyperspectral images as the sequence of information and identified the additional information. They also proposed an activation function and parameter rectified function in the method for analyzing the sequence of data of the image.

R. Ghosh [33] has proposed an RNN based approach for the offline verification and recognition of individual signatures. The working methodology of his approach was divided into three steps as preprocessing, feature extraction, and classification. In the feature extraction step, numerous essential and maneuvering features were mined and the created feature vector was trained by LSTM and bidirectional LSTM approaches. For assessing the accuracy of the approach six different public signature datasets such as GPDS-300, GPDS synthetic, CEDAR, MCYT-75, BHSig260 Bengali, and BHSig260 Hindi were considered and the result was superior in comparison to other modern approaches. M. E. Paoletti et al., [34] also worked on hyperspectral image classification based on scalable RNN architecture. S. Gheisari et al., [35] have proposed a combination of CNN (VGG-16, ResNet-50) and RNN based approaches for the detection of glaucoma from fundus images. Their proposed approach extracts spatial features and temporal features from fundus image and fundus video respectively. A Collection of 1810 images and 295 fundus videos were trained and tested on CNN and their proposed method (CNN+RNN) and the result stated that the CNN+RNN method was superior. K. Kobayashi et al., [36] have proposed an approach called CRNN, a combination of CNN+RNN+FC based automated methods for detecting the behavior of mouse scratching. Scratching is an experimental behavior like itching or psychological stress of an animal. The dataset was trained and tested on CRNN and based on the performance of the model they concluded that their method could be used to study disease.

4. Long Short Term Memory (LSTM)

LSTM is an advanced RNN and is considered as a solution to the vanishing gradient problem of RNN. LSTM uses a memory unit consists of three parts- the first part is called a forget gate which decides whether the information coming from the previous timestamp needs to be remembered or is irrelevant and can be forgotten, the second part is called as input gate which manages the flow of information into the memory and the last part is known as the output gate which passes the updated information from the current timestamp to the next timestamp. The forget gate's operation mechanism is like a one layer neural network and its activation state can be calculated as:

$$F_t = \sigma\left(X[Y_t, O_{t-1}, M_{t-1}] + b_f\right) \tag{4}$$

Where, Y_t is the input vector, O_{t-1} and M_{t-1} are output and memory of previous LSTM cell respectively. X and b_f are the weight and bias vector and σ is sigmoid function. A neural network with tangent function as well as activation values from the previous memory cells are used to calculate the the current memory. These calculations can be done by the following equations.

$$I_t = \sigma\left(X[Y_t, O_{t-1}, M_{t-1}] + b_i\right) \tag{5}$$

$$M_t = F_t.M_{t-1} + I_t.tanh\left([Y_t, O_{t-1}, M_{t-1}]\right) + b_c \tag{6}$$

The received data and information are carried to the output gate and the output computations are as follow:

$$\sigma_t = \sigma\left(X[Y_t, O_{t-1}, M_{t-1}] + b_0\right) \tag{7}$$

$$O_t = o_t.tanh(M_t) \tag{8}$$

LSTM was mainly used for speech recognition. Later, some researchers have used LSTM for image classification also. A. A. Nahid et al., [37] have proposed biomedical breast cancer image classification methods based on LSTM, CNN, and a combination of CNN and LSTM deep learning architecture. They have used the BreakHis breast cancer image dataset consisting of four groups of images such as 40X, 100X 200X, and 400X. The goal of the proposed models was to categorize images into 'benign' and 'malignant' classes. After feature extraction, SVM and Softmax classifiers were also used for making decisions. In their study, the CNN model obtained the best accuracy in comparison to LSTM and CNN+LSTM models. Y. Zhou et al., [38] presented an LSTM based technique to analyse multiple variable time series data. The method classified the crops of a particular region by using the multi-temporal SAR data and the optical images with high resolution. ZY-3 images and Sentinel-1A SAR datasets were considered for classifying the types of crops of two regions of China. The result of classification reflected that the proposed model provided 5 percent more overall accuracy in comparison to other traditional methods. Motivated from human visual perception and visual memorial machinery Y. Zhao et al., [39] proposed a visual LSTM (VLSTM) based combined approach for the classification of fabric defects. The approach consists of 3 segments such as-1) VP-based segment implemented by SCAE, 2) VSTM based segment implemented by shallow CCN and 3) VLTM based segment implemented by a non-local neural network. Three different datasets such as 'DHU-FD-500', 'DHU-FD-1000', and 'Aliyun-FD-10500' were considered for performance evaluation and the result reflected that the proposed method was superior to other traditional and deep learning-based models. D. Li et al., [40] have proposed a 'Multi-Instance Learning' (MIL) technique based on LSTM for classifying images of Chinese paintings. Their work was divided into three phases- firstly, designing a system for modeling multiple instances which converts the painting images to multi-instance bags. Secondly, designing an order generator for constructing a set of discriminating instances and transforms the multiple-instances bags into a well-organized and equivalent length of semantic order. Finally, the design of the LSTM approach with responsiveness mechanisms for analyzing and performing the semantic and discriminative set of instances. The assessment showed that the developed technique was better than other approaches.

J. Amin et al., [41] proposed an LSTM based automated approach to classify brain tumors from MRIs. They used BRATS and SISS-ISLES datasets in their experiment and achieved 98% accuracy. The result proved that the method accurately classified high and low-grade tumors and it could be helpful for radiologists for classifying brain tumors correctly. F. Demir [42] also proposed a deep LSTM based automated method for detecting COVID-19 patients from the X-ray image of chest. The method comprised of 2 steps, pre-processing and deep LSTM model. In the first step, to improve the classification accuracy gradient operator was considered. The deep LSTM model performed classification consisting of 5 layers such as LSTM, FC layer, ReLU, dropout, and softmax layer. The approach achieved 100% accuracy on 80% train- 20% test of the dataset. The considered dataset was consisting of a total number of 1061 CX- images comprising of 'Normal', 'COVID-19' and 'Pneumonia'.

5. Generative Adversarial Network (GAN)

GAN, an unsupervised deep learning approach is a combination of a Generator model 'G' and a Discriminator model 'G'. The G generates the image and the D determines the correctness of the generated image. In GAN, the two models compete against each other in a min-max game. During the process, the goal of G is to generate more realistic images so that it can fool the D. Similarly, the goal of D is identifying the real image from the fake images generated by G. This process goes on until the output of G becomes close to the actual input. The value function of D and G can be represented as follows:

$$min_G max_D V(D, G) = E_{s_i \epsilon \ p_{real}(s)}[logD(s_i)] + E_{z \epsilon \ p_f ake(z)}$$

[log(1-D(G(z)))] (9)

Where, s_i is a sample image, p_{real} and p_{fake} are real and fake images respectively, and a noise vector is represented by z. The goal of GAN is to map p_{fake} to p_{real}. G accepts z as input and produces fake images $G(z)$. The G is trained to minimize $log(1 - D(G(z)))$. and D is to maximize the probability of allocation correct classes to the training image, that is to maximize $logD(s_i)$ [43].

Applications of GAN are image processing, speech and audio processing, medical information processing, image classification, etc. Q. Kong et al., [43] worked on a GAN based image classification model called AC-GAN. Their model consists of three components- firstly, a classifier, secondly, a conditional GAN which generates a labeled sample, and thirdly, a Gaussian multi-layer perceptron. For each generated sample compensation was conceived for measuring the amount of information. MNIST, CIFAR-10, Fashion-MNIST, and Tiny-ImageNet datasets were considered for evaluating the performance of the developed technique. Many researchers have been working on the development of hyperspectral image classification methods based on GAN or a variation of GAN with a combination of other approaches. For example, Z. He et al., [44] and Y. Zhan et al., [45] have worked on a GAN-based semi-supervised method for classifying hyperspectral images. Z. Zhong et al., [46] worked on GAN-CRF-based structure and H. Gao et al., [47] have also proposed a multi-discriminator GAN (MDGAN) based approach for classifying hyperspectral images. C. Tao et al., [48] have suggested a variational GAN-based semi-supervised hyperspectral image classification method to overcome the challenge of a very less number of labeled samples, and F. Zhang et al., [49] suggested an optimized training method. The suggested method was a combination of Progressive Growing GAN (PG-GAN) and Wasserstein GAN Gradient Penalty (WGAN-GP).

J. Feng et al., [50] presented a CA-GAN (Collaborative learning + Attention mechanism-based GAN) for solving problems of small sample size in the hyperspectral image (HSI) classification. H. Liang et al., [51] have also worked on a GAN-based approach for the same. They have developed a GAN-based semi-supervised technique called adaptive weighting feature fusion (AWF^2-GAN) for HSI classification and to overcome the challenges of redundant spectral information, weak spectral-spatial representation and lesser labeled samples faced during classification. 'Indian Pines' (IN) and 'Pavia University' (UP) datasets were assumed by the AWF^2-GAN method for evaluating its performance and the results were matched with other techniques like- HS-GAN, 3D-GAN, SS-GAN and AD-GAN. The comparison results stated that the suggested AWF^2-GAN technique was superior to others. Like other deep learning models, the GAN model is also been used for detecting COVID-19 disease from CR- image. S. Motamed et al., [52] have proposed a randomized GAN (RANDGAN) technique intending to detect COVID-19 disease from the images of chest X-ray (CR). Z. Ji et al., [53] have worked on a GAN-based method for the classification of zero-shot images. Zero-shot image classification is a process of classifying an unseen image (unlabeled image) from the description of the unseen image and the knowledge obtained from the seen image. For the classification of zero-shot images, the researchers have proposed a method called triple discriminator generative adversarial network (TDGAN). TDGAN consisted of dual discriminator GAN and a text reconstruction network. Two datasets, 'Caltech-UCSD Birds-2011' (CUB) and 'North America Birds' (NABirds) were considered to measure the efficacy of TDGAN. The efficacy of the TDGAN technique was matched with the other eight advanced zero-shot classification techniques and it achieved superior performance.

6. Restricted Boltzmann Machine (RBM)

RBM is an undirected probabilistic graphical model based on a bipartite graph. The model contains visible units $V = \{0,1\}^v$ in one part and a hidden unit $H = \{0,1\}^h$ in other parts of the graph. The visible units receive input data and the hidden units generate feature information. Using these visible and hidden units the joint energy function can be expressed as:

$$E(V,H;\theta) = -\sum_{i=1}^{v} b_i v_i - \sum_{i=1}^{H} a_j h_j - \sum_{i=1}^{v}\sum_{j=1}^{H} w_{ij} v_i h_j \quad (10)$$

$$= -b^T V - a^T H - V^T wH$$

Where, $\theta = \left(b_i, a_j, w_{ij}\right)$. i and j are visible and hidden unit respectively. w_{ij} is weight between i and j, b_i is the bias of the visible unit and a_j is the bias of the hidden unit [54]. The structure of an RBM is shown in Figure 2.

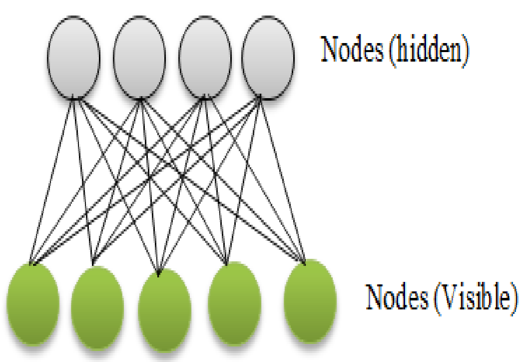

Figure 2. An overview of RBM architecture

Applications of RBM are image classification, dimensionality reduction, feature learning, topic modeling, etc. Here, we have discussed some image classification methods which are developed based on the RBM deep learning approach. N. Lu et al., [55] proposed a method to motor imagery classification based on RBM and Fast Fourier Transformation (FFT). The motor imagery is generally measured by electroencephalography (EEG). EEGs are obtained by Fast Fourier Transformation (FFT) and wavelet package decomposition (WPD). They performed extensive and systematic experiments on the dataset and the result showed improvement over other modern methods. The result of the experiment also generated a few suggestions and concluded that FFT is better than WPD. A. A. Alani [56] has proposed a combination of RBM and CNN based method for recognizing the handwritten digits of Arabic language. In the proposed method, the RBM extracts extremely valuable features from the raw image and they are fed a CNN for categorization. They used 'CMATERDB 3.3.1' dataset for performance evaluation of the method and resulted with 98.59 percent classification accuracy. A. Al Nahid et al., [37] have developed an RBM with a backpropagation-based approach for classifying breast cancer images. They used scaled conjugate gradient techniques for backpropagation and BreakHis dataset for experimental performance evaluation.

M. Amir et al., [57] have developed an approach called NRBM to solve the difficulty of multi-class image classification tasks. The difficulty in multiclass image classification is high dimensional feature space. They developed the model intending to achieve three goals, firstly, reduction of feature dimensionality, secondly, application of Softmax classifier in the final layer of NRBM, and lastly, conduction of adequate experimentations to verify the efficiency of the suggested approach. A total five number of MNIST datasets were considered for the experimental evaluation and the result indicated that the suggested model was better than other models. Like others, K. Tan et al., [58] also worked on RBM-based hyperspectral image (HSI) classification. They proposed an HSI classifier called parallel Gaussian- Bernoulli RBM (GBRBM). The method consists of several GBRBM and extracts features from HSIs. A. M. Mahmoud et al., [59] proposed a 'Contractive Autoencoder' + 'Restricted Boltzmann Machine' based method for the diagnosis of female brain from the functional magnetic resonance imaging (fMRI) scans. The researcher's main objective behind the proposed approach was to investigate the performance of classification of autism spectrum disorder (ASD) and typical control (TC) from fMRI scans of the female brain. ABIDE dataset was considered for performance evaluation and resulted with 88.17 percent classification accuracy. Recently, R. de Souza et al., [60] developed an automated Parkinsons Disease (PD) detection system based on fuzzy optimum path forest and RBMs.

7. Deep Belief Network (DBN)

DBNs are considered as one of the first non-convolutional models. It consists of a stack of RBMs. We know, RBM consists of a visible layer and a hidden layer. As DBN consists of a stack of RBMs, therefore, each RBM's visible layer is connected to the hidden layer of the previous RBM. An overview of DBN architecture is shown in Figure 3. In training, each RBM is trained one after another in an unsupervised manner using a contrastive divergence procedure [61]. The output of each RBM is passed as input to the next RBM. The bottom layer RBM is trained using the input data as a visible unit, next RBM is

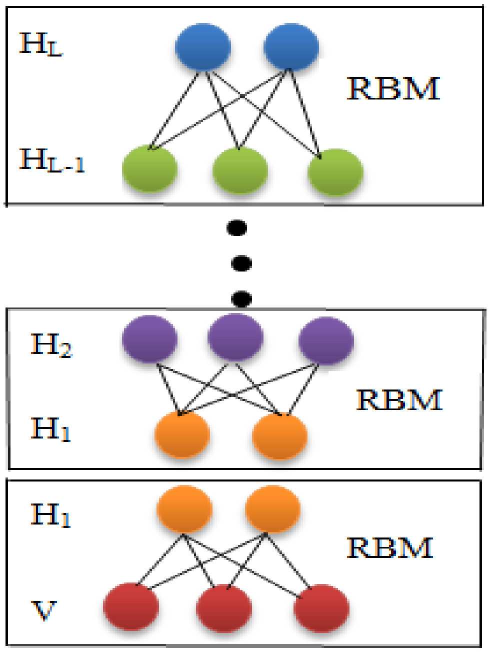

Figure 3. An overview of DBN architecture

trained using the output of the hidden layer (H_1) of RBM1, and so on. The applications of DBN are image classification, speech recognition, audio classification, natural language understanding, etc.

A. Abdel-Zaher et al., [62] presented a DBN-based automatic breast cancer detection system intending to detect cancer at an early stage. Their method is called DBN-NN which could detect breast cancer based on the unsupervised pre-training phase of DBN followed by the supervised back-propagation phase of the neural network. The unsupervised DBN with pre-trained backpropagation achieved better classification accuracy. Wisconsin Breast Cancer dataset (WBCD) was used for assessment of the method to identify misclassified samples, and in comparison to other methods the proposed method outperformed with 99.68 percent classification accuracy. S. A. Pundir et al., [63] proposed a smoke detection system with an aim to early detection of fire. The working procedure of the suggested technique was comprised of four steps. The first step was to determine the moving targeted region. Secondly, detection of smoke color. The third step was to find out the texture and intensity feature of smoke by LECoP computation. Finally, smoke and non-smoke images were classified by DBN. They considered a variety of smoke videos such as wildfire, hill base, indoor and outdoor smoke videos for evaluating the performance of the suggested technique. In contrast to former smoke detection methods, the proposed method attained 99.51 percent classification accuracy. They also concluded that the method could be extended for fire detection. S. Li et al., [64] and H. Chen et al., [17] also worked on a DBN-based method for classifying HSI.

J. Jo et al., [65] have developed a road crack detection system called autonomous crack detection system (ACDS) with a motive to reduce the maintenance cost. The suggested method was comprised of 3 different steps. The first step was called the image acquisition step which collected RGB and infrared images. The infrared images can detect very tiny or invisible road cracks. The second step was called the image processing step, which was comprised of 5 filtering techniques and could be used for removing noise and extracting features from images. The final step was called the classification step where five DBN classifiers analyzed the images for detecting cracks in the images. A. V. Reddy et al., [66] presented a hybrid DBN-based technique to classify glioblastoma (GBM) brain tumors from magnetic resonance imaging (MRI). The developed model integrated discrete wavelet transform (DWT) with DBN for improving efficiency in the form of reducing training time and complexity. It was also a three-stage procedure including preprocessing, dimensionality reduction, and classification. Time complexity, space complexity, and classification accuracy of the developed technique were found higher than former classification techniques. M. Li et al., [67] developed a DBN-based method called DBN-ML for classifying the fine of a pit-mine region of Wuhan city from the Ziyuan-3 imagery. In comparison to other DBN and CNN-based models, performance of the suggested method was better with an overall accuracy of 95.10 percent.

8. Performance Comparison of various Approaches in various Applications

The comparative study depicts the image classification accuracy of different methods in various domains of application especially, HSI classification, breast cancer classification, brain tumor, and covid-19 detection as shown in Table I, II, III, and IV respectively.

From Table I, it is observed that in hyperspectral image classification different methods attained different accuracies with the 'Indian Pines' and 'Pavia University' dataset. However, among all these methods the NLSS-RNN method with Indian Pines and Pavia University dataset attained the highest overall accuracy (OA) of 98.75% and 99.77% respectively. Similarly, from Table II, it is observed that in the area of breast cancer classification the DBN-NN method with 'Wisconsin Breast Cancer Dataset' (WBCD) attained the best 99.68% classification accuracy. However, in the analysis of brain tumor detection with various methods in Table III, the performance of the two methods GoogleNet and LSTM is the same. Both the methods attained 98% accuracy with the 'MRI' dataset. Finally, Table IV contains the performance of three methods for COVID-19 detection with the 'chest X-ray' image dataset. In this case, the deep LSTM method attained 100% classification accuracy. However, it can't be claimed as the best method as the classification accuracy was not evaluated on a challenging dataset.

TABLE I. Result comparison of some hyperspectral image classification methods with Indian Pines (IP) and Pavia University (PU) datasets

Methods	Overall Accuracy in % with IP	Overall Accuracy in % with PU
QGASR-CNN [17]	94.1	91.6
RNN-GRU-PRetanh [28]	88.63	88.85
LSS-RNN [30]	98.36	99.63
NLSS-RNN [30]	98.75	99.77
RNN-SRU [34]	80.48	93.51
Semi-Sup. GAN [44]	75.62	77.94
MDGAN [47]	95.73	94.67
GAN [49]	97.56	98.91
AWF^2-GAN [51]	97.83	98.68
JSSC-DBN [64]	96.29	97.67

TABLE II. Result comparison of Breast cancer classification methods

Methods	Accuracy in %	Dataset Used
CNN with TL+Data Aug. [14]	95.27	BreakHis
CNN [37]	91	BreakHis
RBM [68]	88.7	BreakHis
DBN-NN [62]	99.68	WBCD

9. Summary

This section contains a brief summary of the above-discussed image classification methods which has been developed based on the deep learning approaches. This is summarized in Table V. In Table V, the 'Model Used' column depicts the deep learning approach which has been considered to construct the methods. The key demerits

TABLE III. Result comparison of Brain tumor detection methods

Methods	Accuracy in %	Dataset Used
GoogleNet [12]	98	MRI
VGG-16 [20]	95.71	MRI
LSTM [41]	98	MRI
DWT-PCA-DBN [66]	97	MRI

TABLE IV. Result comparison of Covid-19 detection methods

Methods	Accuracy in %	Dataset Used
CNNs [2]	99.7	Chest X-ray images
Deep LSTM [42]	100	Chest X-ray images
Hybrid (mAlexNet+BiLSTM) [19]	98.70	Chest CT X-ray images

of deep learning approaches are the requirement of large amount of training data and powerful computer. In view of these demerits the researchers may also consider using Few-Shot Learning (FSL) [69], One-Shot Learning (OSL) [70] and Zero-Shot Learning (ZSL) [71] paradigms.

TABLE V. Summary

Cited in	Year	Model Used	Observation	Research Gap
[2]	2021	CNN	They proposed 5 pre-trained CNN models.The model ResNet50 provided the best performance and attained 96.1%, 99.5% and 99.7% accuracy with dataset1, 2 and 3 respectively. It is applied to detect COVID-19.	High computational complexity.
[10]	2016	CNN	The methods improved feature learning automatically for image classification and also provided better classification accuracy in comparison to SVM and ELM. DCNN-5 attained 97.39% and DCNN-7 with 97.91% classification accuracy. It is used to classify tumor images.	Unsatisfactory classification accuracy and can be improved by using AlexNet or VGGNet. Secondly, the usage of a very deep model leads to the problem of overfitting and more computational cost.
[11]	2018	CNN	Utilization of pre-trained DL and TL. Experiments were performed on both augmented images and original images without augmentation. The method obtained 80% and 98.61% accuracy with original images and augmented images respectively. It is applied for classifying skin cancer.	Classification performance may be improved and a large number of labeled images are required.
[12]	2019	CNN	For a dataset with a small number of images the method's performance was good and classified 3 types of tumors with 98% accuracy. It is used for classifying brain tumor.	Poor performance of transfer learning, Misclassification of the meningioma category of brain tumor and the problem of overfitting with a smaller training dataset.
[13]	2020	CNN	For reduction of overfitting on training layers, the researchers used a callback parameter in the model which stopped training after 3 iterations if there was no accuracy improvement.VGG16provided better accuracy than VGG19. It is used to classify lung disease.	Lake of datasets, algorithmic complexity and Performance of the model can be improved.
[14]	2020	CNN	The method improved classification accuracy and computational complexity. It obtained 91% accuracy with the BreakHis dataset. It is applied to classify breast cancer images.	Classification performance can still be improved.
[16]	2020	CNN	The method was aimed to reduce computational time and energy uses.	Although the method sped up CNN-based models there was a degradation of classification accuracy.
[17]	2020	CNN	The proposed method called QGASR-CNN attained 94.1% classification accuracy and solved the problem of salt and pepper misclassification. It is applied to classify remote sensing images.	Less classification accuracy.

Cited in	Year	Model Used	Observation	Research Gap
[21]	2021	CNN	Data augmentation and VGG-16 CNN techniques were used and obtained 95.71% classification accuracy. It is used for classifying fundus images.	Classification accuracy might be improved.
[22]	2021	CNN	The method classified a variety of Corneal Ulcers with 92.73% accuracy. This method is used to detect corneal ucers.	SUSTech-SYSU dataset and the method may not be executed on the general computers of a hospital.
[23]	2021	CNN	Used four CNN models VGG19, ResNet-101, Inception ResNet-V2 and Xception to classify and detect 4-types of paddy leaf diseases and attained 81.43%, 91.52%, 92.68% and 89.42% classification accuracy respectively. It is used for detecting the disease of paddy leaves.	Among the four CNN models, the maximum accuracy obtained was 92.68% only and hence the performance needs to be improved.
[25]	2021	CNN + LSTM	They used CNN for detecting fundus images and LSTM for generating descriptions of those images. The method achieved 90% accuracy. This method is applied for diabetic retinopathy detection.	Performance is less, 90% only.
[28]	2017	RNN	They used RNN first time for HSI classification and proposed PRetanh activation function which performed better than tanh. The method provides better performance than SVM-RBF and CNN.	The proposed method provided 86.33% average accuracy and 88.85% overall accuracy which can be improved.
[29]	2018	RNN	The method was for the classification of multi-label images. The performance of the method was better than other methods such as CNN-softmax, WARP, ResNet-50, etc.	Achieved different accuracies with different datasets like NUS-WIDE, MS-COCO and ESP Game respectively. However, the observed performance was not outstanding.
[30]	2018	RNN	Proposed two approaches LSS-RNN and NLSS-RNN. They were tested on three HSI datasets as, ‘Indian pines’, ‘University of Pavia’, and ‘Salinas’ respectively. The LSS-RNN achieved 98.36%, 99.63% and 97.11% overall accuracy with these three datasets respectively and NLSS-RNN achieved 98.75%, 99.77% and 97.23% overall accuracy. . It is used for classifying hyperspectral images.	RNN framework can be improved.
[31]	2019	RNN	They developed a multi-label image classification approach. They used Pascal VOC, MS-COCO, and BNID datasets for evaluating performance and achieved mAP of 85.6%, 64.72% and 53.48% respectively for each dataset. It is used for the classification of multi-label images.	Performance need to be improved.

Cited in	Year	Model Used	Observation	Research Gap
[33]	2020	RNN	The method could recognize and verify an individual's offline signature. The performance was evaluated on six different signature datasets such as 'GPDS synthetic', 'GPDS-300', 'MCYT-75', 'CEDAR', 'BHSig260 Hindi', and 'BHSig260 Bengali' and achieved 96.08%, 98.02%, 99.39%, 99.94%, 99.28% and 99.37% recognition accuracy (RA) respectively.	
[34]	2020	RNN	The proposed method achieved outstanding classification accuracy and computational performance with large datasets. This method is applied for HSI classification.	The method could not solve the problem of overfitting.
[35]	2021	CNN + RNN	The method extracted both spatial and temporal features from fundus images and videos respectively. It achieved 95% sensitivity and 96% specificity in recognizing glaucoma.	Firstly, the dataset's size was comparatively smaller and from a uniform population. Secondly, two different clinicians labeled the images and videos. Thirdly, the cameras used for capturing retinal fundus images and videos were different.
[36]	2021	CRNN (CNN + RNN + FC)	The model predicted scratch and non-scratch segments of 81.6% and 87.9% respectively for the test dataset. It is used to detect the behavior of mouse scratching.	The performance can be improved.
[42]	2021	LSTM	The method achieved 100% accuracy with 80% training and 20% testing of the MCWS dataset. The performance also varied with the variation of training-testing of the dataset. It is applied to detect COVID-19 from CX images.	Though, it achieved 100% accuracy it cannot be demanded that it is superior to other models since the performance was not evaluated with a challenging dataset
[37]	2018	LSTM, CNN, CNN + LSTM	They used the BreakHis dataset for evaluating performance and achieved a sensitivity, specificity, recall, F-Measure, and classification accuracy of 93%, 96%, 96%, 93%, and 91% respectively. It is applied to classify breast cancer images.	The classification accuracy is 91% only which needs to be improved and the used dataset was very small.
[38]	2019	LSTM	ZY-3 optical images and multi-temporal Sentinel-1A SAR datasets were used for evaluating the classification performance of crop types of the Hunan and Guizhou regions of China. The method achieved 5% more overall accuracy than SVM and RF.	The attained overall accuracies for Hunan and Guizhou were only 83.71% and 80.71% respectively.
[39]	2019	Visual LSTM	The mean accuracy of the proposed method was 97.64%. It was superior in comparison to other methods like KNN, Wavelet-based BPNN, AlexNet and VGG. It is applied to classify fabric defects.	It was a comparatively complex model in comparison to Wavelet-BPNN.

Cited in	Year	Model Used	Observation	Research Gap
[40]	2020	LSTM	Three datasets were used and 91.5% average accuracy was achieved. The method was superior to other methods like MI-SVM, DD-SVM, MILES, CNN, LSTM, etc. In addition to image classification, the method was also applicable to image surface defect detection.	Accuracy improvement needed.
[41]	2020	LSTM	The method classified high and low grade brain tumors with 98% accuracy. The method performed better comparatively.	
[43]	2019	GAN	F-Score of the model on CIFAR-10, MNIST, F-MNIST, and Tiny-ImageNet datasets were 86.3, 96.5, 89.0, and 59.4 respectively. The F-score of the datasets demonstrated that it could generate informative labeled images which were sure to be effective to improve the classification accuracy.	The result needs to be improved.
[46]	2019	GAN + CRF	The method solved three problems of HIS classification- large spectral features of training data, less number of labeled samples, and the noisy maps of classification. The method achieved 97.61% overall accuracy and was superior to other DL-based methods.	
[47]	2019	GAN	The method used Indian Pines, Pavia University and Salinas datasets for evaluating the performance and achieved 97.68%, 98.50% and 97.74% average accuracy respectively. The result also reflected that the method MDGANs was superior to KNN, NN, SVM and CNN methods. It is used for classifying hyperspectral images.	The extraction of spatial features can be improved.
[48]	2020	GAN	The method solved the problem of HSI classification where the labeled samples were very few.	Existence of overfitting and poor discriminant feature space.
[51]	2021	GAN	The method solved the challenges of redundant spectral information, less labeled samples, and spectral-spatial information for HSI classification. The method was evaluated on Indian pines and Pavia University datasets and achieved a very good classification accuracy of about 96% and 98% respectively. The training time of the proposed method was 5-7 times lesser than SS-GAN and AD-GAN methods.	

Cited in	Year	Model Used	Observation	Research Gap
[52]	2021	GAN	The proposed method called RANDGAN solves the problem of high labeled data requirement tasks for all classes. Even for training, it does not require any COVID-19 CX-ray images. The model achieved 77% AUC.	The performance needs to be improved.
[53]	2021	GAN	A GAN-based method called TDGAN used for zero-shot image classification. TDGAN is better than other methods and achieved 44.2% and 36.7% classification accuracy on CUB and NA Birds dataset respectively.	Classification performance is not satisfactory.
[56]	2017	RBM + CNN	The method was superior to CNN, RBM-SVM and Sparse RBM-SVM method and obtained 98.59% accuracy on CMATERDB 3.3.1dataset. The method is applied to detect Arabic handwritten digits.	
[68]	2018	RBN	BreakHis dataset was used for the evaluation of the method. The dataset was divided into four groups of magnification factor(X) as 40X, 100X, 200X and 400X and acquired 88.7%, 85.3%, 88.6% and 88.4% accuracy respectively. Used for classifying breast images.	The classification performance was not outstanding and can be improved by RBM+ CNN.
[57]	2019	RBM	The proposed method is called NRBM and its execution time, training and testing performances on the MNIST dataset are better than DBN, GBN and RBM-based methods.	
[58]	2019	RBM	The method is called Gaussian Bernoulli RBM (GBRBM) which extracted features from HSI and reduces prediction time. The Salinas obtained 95.94% and the Pavia University dataset obtained 96.22% accuracy. GBRBM's performance was better than SVM, DBN and Gcforest methods.	
[59]	2020	RBM + Contractive Autoencoder	The method obtained 88.17% classification accuracy and it was superior to SVM, KNN, and RF classifiers. This method is applied to diagnose the female brain.	The classification accuracy was 88.17% which is not considered an outstanding performance; hence it needs to be improved.
[60]	2021	RBM	With the different number of features, the model's performance also varies. The overall performance of the model is superior to other models like OPF, KNN and SVM. It is applied to identify Parkinsons disease.	The model takes a longer training time than OPF, KNN and SVM.
[62]	2016	DBN	The technique was tested on the WBCD dataset and obtained 99.68% classification accuracy. It's used for classifying breast cancer.	

Cited in	Year	Model Used	Observation	Research Gap
[63]	2017	DBN	4 types of datasets were used for evaluating the performance of the method and achieved 99.51% classification accuracy. The result of this method was superior to ANN and autoEncoder. This method is applied to detect smoke.	Though, its accuracy was very good but, the method took more time for evaluation.
[64]	2019	DBN	The method used Indian pines and Pavia University dataset for evaluating the performance and achieved 97.7% and 95.8% accuracy.	The computation time of the method was higher and classification accuracy also could be improved.
[72]	2020	DBN	The methods (FRDBN, PRPDBN) obtained approximately 99% accuracy and were considered novel methods for HSI classification.	The training time of the methods was high.
[65]	2020	DBN	The method (ACDS) was developed for detecting cracks on road surfaces. It obtained 90% precision with RGB images and 92% precision accuracy with RGB+ Infrared images. ACDS was better than SVM, Boosting, and ConvNets method.	The highest precision accuracy was 92% only, therefore, precision accuracy could be improved.
[66]	2020	DBN	The method DWT-PCA-DBN obtained 97% overall accuracy. Its performance was better than non-deep learning methods as well as DBN and DWT-DBN DL methods. Initialization of weight and epochs were important factors. . It is applied for tumor detection.	A large number of iterations might produce a better result but it degrades performances by the means of high training time and overfitting.
[67]	2021	DBN	The proposed method is called DBN-ML. It achieved 95.10% overall accuracy and its result is superior to FS-SVM, DBN-S, DBN-RF, DBN-SVM, CNN and DCNN models. This method is used to classify geo-environments.	The time complexity of DBN-ML is higher. Its accuracy can also be improved.

10. Conclusions and Future Work

Deep learning is widely used in image classification domain. Applications of image classification techniques can be in several areas such as cancer detection, tumor detection, road crack identification, signature identification, hyperspectral image (HSI) analysis, etc. Although CNN is considered as a standard deep learning approach for image classification, several other approaches are also considered for developing image classification methods. In our survey on deep learning approach-based image classification methods, it is found that different methods have been developed for a specific image classification problem. However, the DL based methods like DCNN-5, QGASR-CNN, CRNN, TDGAN, etc. can't deliver the highest level of desired classification accuracy. Issues like overfitting, higher time complexity and the need of large amount of labeled images are some of the problems encountered on those methods. Further research and investigation are essential in order to overcome those limitations. Researchers may adopt Few-Shot Learning, One-Shot Learning, and Zero-Shot Learning paradigms to overcome the limitations of deep learning.

Conflicts of Interest

There is no conflicts of interest.

References

[1] R. A. Hazarika, A. Abraham, S. N. Sur, A. K. Maji, and D. Kandar, "Different techniques for alzheimer's disease classification using brain images: a study," *International Journal of Multimedia Information Retrieval*, pp. 1–20, 2021.

[2] A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks," *Pattern Analysis and Applications*, pp. 1–14, 2021.

[3] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural computation*, vol. 1, no. 4, pp. 541–551, 1989.

[4] Y. LeCun, L. D. Jackel, L. Bottou, C. Cortes, J. S. Denker, H. Drucker, I. Guyon, U. A. Muller, E. Sackinger, P. Simard *et al.*, "Learning algorithms for classification: A comparison on handwritten digit recognition," *Neural networks: the statistical mechanics perspective*, vol. 261, no. 276, p. 2, 1995.

[5] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, pp. 1097–1105, 2012.

[6] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[7] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.

[8] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

[9] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural networks," *Artificial Intelligence Review*, vol. 53, no. 8, pp. 5455–5516, 2020.

[10] X. Yang, S. Y. Yeo, J. M. Hong, S. T. Wong, W. T. Tang, Z. Z. Wu, G. Lee, S. Chen, V. Ding, B. Pang *et al.*, "A deep learning approach for tumor tissue image classification," *IASTED Biomedical Engineering*, 2016.

[11] K. M. Hosny, M. A. Kassem, and M. M. Foaud, "Skin cancer classification using deep learning and transfer learning," in *2018 9th Cairo international biomedical engineering conference (CIBEC)*. IEEE, 2018, pp. 90–93.

[12] S. Deepak and P. Ameer, "Brain tumor classification using deep cnn features via transfer learning," *Computers in biology and medicine*, vol. 111, p. 103345, 2019.

[13] H. Moujahid, B. Cherradi, O. El Gannour, L. Bahatti, O. Terrada, and S. Hamida, "Convolutional neural network based classification of patients with pneumonia using x-ray lung images," *Transfer*, vol. 2, no. 99.41, p. 16, 2020.

[14] P. Yamlome, A. D. Akwaboah, A. Marz, and M. Deo, "Convolutional neural network based breast cancer histopathology image classification," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2020, pp. 1144–1147.

[15] S. Srinivasan, V. Ravi, V. Sowmya, M. Krichen, D. B. Noureddine, S. Anivilla, and K. Soman, "Deep convolutional neural network based image spam classification," in *2020 6th Conference on data science and machine learning applications (CDMA)*. IEEE, 2020, pp. 112–117.

[16] Z. Chen, J. Zhang, R. Ding, and D. Marculescu, "Vip: Virtual pooling for accelerating cnn-based image classification and object detection," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2020, pp. 1180–1189.

[17] H. Chen, F. Miao, and X. Shen, "Hyperspectral remote sensing image classification with cnn based on quantum genetic-optimized sparse representation," *IEEE Access*, vol. 8, pp. 99 900–99 909, 2020.

[18] F. Zhang, M. Yan, C. Hu, J. Ni, and F. Ma, "Integrating global spatial features in cnn based hyperspectral/sar imagery classification," *arXiv preprint arXiv:2006.00234*, 2020.

[19] M. F. Aslan, M. F. Unlersen, K. Sabanci, and A. Durdu, "Cnn-based transfer learning–bilstm network: A novel approach for covid-19 infection detection," *Applied Soft Computing*, vol. 98, p. 106912, 2021.

[20] V. K. Waghmare and M. H. Kolekar, "Convolutional neural network-based automatic brain tumor detection," in *Evolving Technologies for Computing, Communication and Smart World*. Springer, 2021, pp. 463–474.

[21] A. Anitha, P. Padmapriya, P. Preethi, T. Swetha, and K. Banumathi, "Fundus image classification of eye disease using cnn method (convolution neural network)," 2021.

[22] J. Gross, J. Breitenbach, H. Baumgartl, and R. Buettner, "High-performance detection of corneal ulceration using image classification with convolutional neural networks," in *Proceedings of the*

54th Hawaii International Conference on System Sciences, 2021, p. 3416.

[23] M. Islam, M. Shuvo, N. Rahman, M. Shamsojjaman, S. Hasan, M. Hossain, T. Khatun *et al.*, "An automated convolutional neural network based approach for paddy leaf disease detection," 2021.

[24] N. Kesav and M. Jibukumar, "Efficient and low complex architecture for detection and classification of brain tumor using rcnn with two channel cnn," *Journal of King Saud University-Computer and Information Sciences*, 2021.

[25] R. Amalia, A. Bustamam, and D. Sarwinda, "Detection and description generation of diabetic retinopathy using convolutional neural network and long short-term memory," in *journal of physics: conference series*, vol. 1722, no. 1. IOP Publishing, 2021, p. 012010.

[26] K. El Asnaoui, "Design ensemble deep learning model for pneumonia disease classification," *International Journal of Multimedia Information Retrieval*, vol. 10, no. 1, pp. 55–68, 2021.

[27] K. Cho, B. van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder–decoder for statistical machine translation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Doha, Qatar: Association for Computational Linguistics, Oct. 2014, pp. 1724–1734.

[28] L. Mou, P. Ghamisi, and X. X. Zhu, "Deep recurrent neural networks for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 7, pp. 3639–3655, 2017.

[29] L. Li, S. Wang, S. Jiang, and Q. Huang, "Attentive recurrent neural network for weak-supervised multi-label image classification," in *Proceedings of the 26th ACM international conference on Multimedia*, 2018, pp. 1092–1100.

[30] X. Zhang, Y. Sun, K. Jiang, C. Li, L. Jiao, and H. Zhou, "Spatial sequential recurrent neural network for hyperspectral image classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 11, pp. 4141–4155, 2018.

[31] F. Lyu, Q. Wu, F. Hu, Q. Wu, and M. Tan, "Attend and imagine: Multi-label image classification with visual attention and recurrent neural networks," *IEEE Transactions on Multimedia*, vol. 21, no. 8, pp. 1971–1981, 2019.

[32] R. Venkatesan and S. Prabu, "Hyperspectral image features classification using deep learning recurrent neural networks," *Journal of medical systems*, vol. 43, no. 7, pp. 1–10, 2019.

[33] R. Ghosh, "A recurrent neural network based deep learning model for offline signature verification and recognition system," *Expert Systems with Applications*, vol. 168, p. 114249, 2021.

[34] M. E. Paoletti, J. M. Haut, J. Plaza, and A. Plaza, "Scalable recurrent neural network for hyperspectral image classification," *The Journal of Supercomputing*, vol. 76, no. 11, pp. 8866–8882, 2020.

[35] S. Gheisari, S. Shariflou, J. Phu, P. J. Kennedy, A. Agar, M. Kalloniatis, and S. M. Golzan, "A combined convolutional and recurrent neural network for enhanced glaucoma detection," *Scientific reports*, vol. 11, no. 1, pp. 1–11, 2021.

[36] K. Kobayashi, S. Matsushita, N. Shimizu, S. Masuko, M. Yamamoto, and T. Murata, "Automated detection of mouse scratching behaviour using convolutional recurrent neural network," *Scientific reports*, vol. 11, no. 1, pp. 1–10, 2021.

[37] A.-A. Nahid, M. A. Mehrabi, and Y. Kong, "Histopathological breast cancer image classification by deep neural network techniques guided by local clustering," *BioMed research international*, vol. 2018, 2018.

[38] Y. Zhou, J. Luo, L. Feng, Y. Yang, Y. Chen, and W. Wu, "Long-short-term-memory-based crop classification using high-resolution optical images and multi-temporal sar data," *GIScience & Remote Sensing*, vol. 56, no. 8, pp. 1170–1191, 2019.

[39] Y. Zhao, K. Hao, H. He, X. Tang, and B. Wei, "A visual long-short-term memory based integrated cnn model for fabric defect image classification," *Neurocomputing*, vol. 380, pp. 259–270, 2020.

[40] D. Li and Y. Zhang, "Multi-instance learning algorithm based on lstm for chinese painting image classification," *IEEE Access*, vol. 8, pp. 179 336–179 345, 2020.

[41] J. Amin, M. Sharif, M. Raza, T. Saba, R. Sial, and S. A. Shad, "Brain tumor detection: A long short-term memory (lstm)-based learning model," *Neural Computing and Applications*, vol. 32, no. 20, pp. 15 965–15 973, 2020.

[42] F. Demir, "Deepcoronet: A deep lstm approach for automated detection of covid-19 cases from chest x-ray images," *Applied Soft Computing*, vol. 103, p. 107160, 2021.

[43] Q. Kong, B. Tong, M. Klinkigt, Y. Watanabe, N. Akira, and T. Murakami, "Active generative adversarial network for image classification," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, 2019, pp. 4090–4097.

[44] Z. He, H. Liu, Y. Wang, and J. Hu, "Generative adversarial networks-based semi-supervised learning for hyperspectral image classification," *Remote Sensing*, vol. 9, no. 10, p. 1042, 2017.

[45] Y. Zhan, D. Hu, Y. Wang, and X. Yu, "Semisupervised hyperspectral image classification based on generative adversarial networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 2, pp. 212–216, 2017.

[46] Z. Zhong, J. Li, D. A. Clausi, and A. Wong, "Generative adversarial networks and conditional random fields for hyperspectral image classification," *IEEE transactions on cybernetics*, vol. 50, no. 7, pp. 3318–3329, 2019.

[47] H. Gao, D. Yao, M. Wang, C. Li, H. Liu, Z. Hua, and J. Wang, "A hyperspectral image classification method based on multi-discriminator generative adversarial networks," *Sensors*, vol. 19, no. 15, p. 3269, 2019.

[48] H. Wang, C. Tao, J. Qi, H. Li, and Y. Tang, "Semi-supervised variational generative adversarial networks for hyperspectral image classification," in *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2019, pp. 9792–9794.

[49] F. Zhang, J. Bai, J. Zhang, Z. Xiao, and C. Pei, "An optimized training method for gan-based hyperspectral image classification," *IEEE Geoscience and Remote Sensing Letters*, 2020.

[50] J. Feng, X. Feng, J. Chen, X. Cao, X. Zhang, L. Jiao, and T. Yu,

"Generative adversarial networks based on collaborative learning and attention mechanism for hyperspectral image classification," *Remote Sensing*, vol. 12, no. 7, p. 1149, 2020.

[51] H. Liang, W. Bao, and X. Shen, "Adaptive weighting feature fusion approach based on generative adversarial network for hyperspectral image classification," *Remote Sensing*, vol. 13, no. 2, p. 198, 2021.

[52] S. Motamed, P. Rogalla, and F. Khalvati, "Randgan: randomized generative adversarial network for detection of covid-19 in chest x-ray," *Scientific Reports*, vol. 11, no. 1, pp. 1–10, 2021.

[53] Z. Ji, J. Yan, Q. Wang, Y. Pang, and X. Li, "Triple discriminator generative adversarial network for zero-shot image classification," *Science China Information Sciences*, vol. 64, no. 2, pp. 1–14, 2021.

[54] S. Li, W. Song, L. Fang, Y. Chen, P. Ghamisi, and J. A. Benediktsson, "Deep learning for hyperspectral image classification: An overview," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 9, pp. 6690–6709, 2019.

[55] N. Lu, T. Li, X. Ren, and H. Miao, "A deep learning scheme for motor imagery classification based on restricted boltzmann machines," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 25, no. 6, pp. 566–576, 2016.

[56] A. A. Alani, "Arabic handwritten digit recognition based on restricted boltzmann machine and convolutional neural networks," *Information*, vol. 8, no. 4, p. 142, 2017.

[57] M. Aamir, F. Wahid, H. Mahdin, and N. M. Nawi, "An efficient normalized restricted boltzmann machine for solving multiclass classification problems," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 8, pp. 416–426, 2019.

[58] K. Tan, F. Wu, Q. Du, P. Du, and Y. Chen, "A parallel gaussian–bernoulli restricted boltzmann machine for mining area classification with hyperspectral imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 2, pp. 627–636, 2019.

[59] A. M. Mahmoud, F. Alrowais, and H. Karamti, "A hybrid deep contractive autoencoder and restricted boltzmann machine approach to differentiate representation of female brain disorder," *Procedia Computer Science*, vol. 176, pp. 1033–1042, 2020.

[60] R. W. de Souza, D. S. Silva, L. A. Passos, M. Roder, M. C. Santana, P. R. Pinheiro, and V. H. C. de Albuquerque, "Computer-assisted parkinson's disease diagnosis using fuzzy optimum-path forest and restricted boltzmann machines," *Computers in Biology and Medicine*, vol. 131, p. 104260, 2021.

[61] G. E. Hinton, "Training products of experts by minimizing contrastive divergence," *Neural computation*, vol. 14, no. 8, pp. 1771–1800, 2002.

[62] A. M. Abdel-Zaher and A. M. Eldeib, "Breast cancer classification using deep belief networks," *Expert Systems with Applications*, vol. 46, pp. 139–144, 2016.

[63] A. S. Pundir and B. Raman, "Deep belief network for smoke detection," *Fire technology*, vol. 53, no. 6, pp. 1943–1960, 2017.

[64] C. Li, Y. Wang, X. Zhang, H. Gao, Y. Yang, and J. Wang, "Deep belief network for spectral–spatial classification of hyperspectral remote sensor data," *Sensors*, vol. 19, no. 1, p. 204, 2019.

[65] J. Jo and Z. Jadidi, "A high precision crack classification system using multi-layered image processing and deep belief learning," *Structure and Infrastructure Engineering*, vol. 16, no. 2, pp. 297–305, 2020.

[66] A. V. Reddy, C. P. Krishna, P. K. Mallick, S. K. Satapathy, P. Tiwari, M. Zymbler, and S. Kumar, "Analyzing mri scans to detect glioblastoma tumor using hybrid deep belief networks," *Journal of Big Data*, vol. 7, pp. 1–17, 2020.

[67] M. Li, Z. Tang, W. Tong, X. Li, W. Chen, and L. Wang, "A multi-level output-based dbn model for fine classification of complex geo-environments area using ziyuan-3 tms imagery," *Sensors*, vol. 21, no. 6, p. 2089, 2021.

[68] A.-A. Nahid, A. Mikaelian, and Y. Kong, "Histopathological breast-image classification with restricted boltzmann machine along with backpropagation," *Biomedical Research*, vol. 29, no. 10, pp. 2068–2077, 2018.

[69] Y. Wang, Q. Yao, J. T. Kwok, and L. M. Ni, "Generalizing from a few examples: A survey on few-shot learning," *ACM Computing Surveys (CSUR)*, vol. 53, no. 3, pp. 1–34, 2020.

[70] N. O'Mahony, S. Campbell, A. Carvalho, L. Krpalkova, G. V. Hernandez, S. Harapanahalli, D. Riordan, and J. Walsh, "One-shot learning for custom identification tasks; a review," *Procedia Manufacturing*, vol. 38, pp. 186–193, 2019.

[71] W. Wang, V. W. Zheng, H. Yu, and C. Miao, "A survey of zero-shot learning: Settings, methods, and applications," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 10, no. 2, pp. 1–37, 2019.

[72] C. Chen, Y. Ma, and G. Ren, "Hyperspectral classification using deep belief networks based on conjugate gradient update and pixel-centric spectral block features," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 4060–4069, 2020.

Nayan Kumar Sarkar (Research Scholar) received his M.Sc. degree in Information Technology from Gauhati University, Guwahati, India, in 2013 and the M.Tech. degree in Information Technology from the Tezpur University, Napaam, Sonitpur, India. He is currently pursuing his Ph.D. degree in Computer Science and Engineering from North Eastern Regional Institute of Science and Technology (NERIST), Arunachal Pradesh, India. His research area of interest includes machine learning, deep learning and image classification.

Moirangthem Marjit Singh (Associate Professor) received the B.Tech. degree in Computer Science and Engineering from North-Eastern Hill University, Shillong, India, in 2001, the M.Tech. degree in Computer Science and Engineering from the North Eastern Regional Institute of Science and Technology (NERIST), India, in 2010, and the Ph.D.(Eng.) degree in Computer Science and Engineering from the University of Kalyani, India, in 2017. He was the Head of the Department of Computer Science and Engineering, NERIST during 12 October 2018 to 25 July 2022. He was the Honorary Joint Secretary of the Institution of Engineers, Arunachal Pradesh State Centre, India during 2019-2021. He also secured First Position in X and Second Position in XII Examinations conducted by CBSE, New Delhi, India, amongst the candidates sent up from Jawahar Navodaya Vidyalayas (JNVs) of North Eastern region states of India, in 1995 and 1997, respectively. He has more than 18 years of teaching and research experience. His research interests include mobile adhoc networks, wireless sensor networks, network security, AI, machine learning, and deep learning. He is a Fellow of IETE New Delhi, India and Fellow of the Institutions of Engineers (India). He is also the senior member IEEE, USA. He was awarded the IE(I) Young Engineers Award 2014–2015 from the Computer Engineering Division, Institution of Engineers, India, and received the Best Paper Award at the ICACCT 2016 International Conference (published by Springer). He was awarded the Gold Medal for getting top position in the M.Tech. program at NERIST, India. He has been one of the reviewers to many reputed journals.

Utpal Nandi received his M.Sc. in Computer Science from Vidyasagar University, Midnapore, Paschim Medinipur, West Bengal, India in 2006. He did his M.Tech in Computer Science Engineering in 2009 and Ph.D. in 2018 from University of Kalyani, Nadia, West Bengal, India. He is currently working as an Assistant Professor in the Dept. of Computer Science, Vidyasagar University, Midnapore, Paschim Medinipur, West Bengal, India. He has more than 10 years of teaching and research experiences. His research interest includes data and image compression techniques, image processing and multimedia technology. He has published more than 26 papers in international journals, book chapters and conferences.