



Ensemble Learning Technique for Cloud Classification

Aarti Kumthekar, Ramachandra Reddy G

Abstract: Automatic cloud classification is one of the important areas of remote sensing for metrological applications. Machine learning and deep learning techniques have been used for automatic classification of the cloud type. Several pretrained models are developed using convolutional neural network (CNN), which is part of deep learning. The classification performance of pretrained networks can be further improved using ensemble methods. Ensemble learning can perform better than single learner. In this paper, we proposed two different ensemble learning techniques: ensemble of CNN and ensemble of classifier. In first approach, CNN ensemble is performed, where the features extracted by two or more CNN are combined together using single classifier. The second method is to ensemble the predictions of different classifiers produced by a single or multiple CNN. The accuracy of cloud classification of the proposed methods has improved compared to without ensemble of pretrained networks.

Keywords: Pretrained network, Cloud classification, Ensemble learning

I. INTRODUCTION

Remote sensing image processing is an important research area with real-life applications beneficial to the society. Remote sensing satellite images contain lot of information that helps in monitoring and modeling the processes on the Earth's surface and their interaction with the atmosphere. It can measure and estimate biological, geographical and physical variables to identify materials on the land cover. But the cloud covers the land surface area and hence becomes difficult to observe the earth surface. The cloud cover area is approximately 50% [1]. Its presence affects the real time image processing and obstructs the services to be provided. Most of the satellite images contain cloud which degrades the quality of image and hinders their application, hence processing of cloud and its classification is a crucial and challenging step.

Many algorithms for cloud classification have been proposed which can be divided into three broad classes as threshold based [2], [3], [4], statistical based [5], [6], [7] and

neural network based [8], [9]. Neural network has better performance as compared to traditional techniques. Nowadays machine learning and deep learning have gained

lot of attention. These techniques deal with huge dataset, acquiring the information about data by getting trained, and based on the learned knowledge it helps to give results of the test dataset. Deep learning techniques allow to build complex algorithms in an easier way. Convolutional neural network (CNN) is one of the powerful techniques which has remarkably used in various applications such as image classification, image segmentation, video recognition, etc. CNN technique was invented in the year 1980, but it came into prominence in the year 2000 with fast computations on Graphics processing Units (GPU). GPU is a single chip processor that performs fast complex mathematical computations. Convolutional Neural Networks (CNNs) have achieved breakthrough in many challenges and domains for image classification or object detection. CNN is a multiple layer's network which is trained to learn the low-level and high-level features of images. But the major problem is finding the appropriate annotated dataset. Hence pretrained networks are developed which are already trained on millions of human annotated general images. Alexnet [10] was the first developed technique by Krizhevsky which gained high popularity because of its accuracy. Basically, it has multiple convolutional and max-pooling layers which extract the feature maps of the image that are followed by fully connected layer for image classification. With several modifications in the CNN architecture, many networks came into existence such as, ZFNet [11], GoogleNet [12], VGG [13], ResNet [14], DenseNet [15], etc. These networks are winner of ImageNet Large Scale Visual Recognition Challenge. They are trained on ImageNet dataset containing millions of human annotated images intended to provide a platform for development of computer vision research. CNN architectures outperform with large properly annotated dataset but have high nonlinearity and show high variance. So, ensemble learning helps in reducing the performance error and reduce the variance by getting predictions from multiple models. Ensemble learning is a new emerging field of machine learning [16] and deep learning [10],[17] which uses different algorithms for training the dataset and choosing the final prediction based on ensembling the predictions of classifiers. Landsat 8 satellite is one of the popular satellites collecting high resolution data of the land surface. Landsat 8 satellite images are used to develop the dataset which is downloaded from USGS Earth Explorer (<https://earthexplorer.usgs.gov/>). Fig. 1 shows the true color multispectral image of Landsat 8 satellite with resolution of 30-meter.

Revised Manuscript Received on December 30, 2019.

* Correspondence Author

Aarti Kumthekar, Department of Communication, School of Electronics Engineering (SENSE), Vellore Institute of Technology, Vellore, India,

Dr. Ramachandra Reddy G*, Senior Professor, Department of Communication, School of Electronics Engineering (SENSE), Vellore Institute of Technology, Vellore, India,

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](http://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

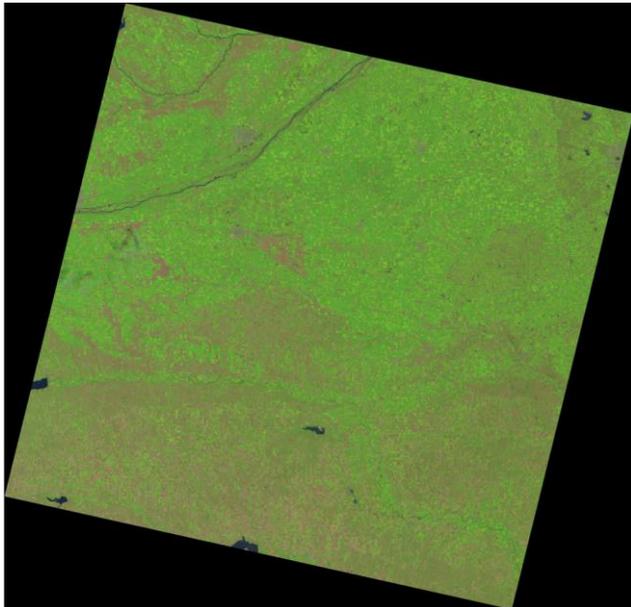


Fig. 1 Landsat 8 Satellite Image

II. PROPOSED METHODOLOGY

The main aim of the proposed methodology is cloud classification using pretrained network which acts as arbitrary feature extractor and classifier for classifying the image. The pretrained networks extract the features using convolutional layer. Different types of pretrained networks are trained with different combinations of convolutional layers to give diversity in features. This simple variation in structure and also variation in hyperparameters like batch size, learning rate, number of epochs and number of hidden layers affect the performance of CNN. The ensemble learning considers the robustness of individual pretrained network and reduces the chances of misclassification. The pretrained networks extract the useful features and these features are concatenated and then classified using classifiers. Apart from the multiple pretrained networks, the multiple classifiers can also be ensemble. Multiple classifiers can have better performance by classifying diverse features obtained from pretrained networks.

In this paper, the first stage deals with ensembles of the pretrained networks such as VGG16, ResNet and DenseNet and extracts the diverse useful features. These features are classified by combining the multiple classifiers such as support vector machine and K-nearest neighbors. The ensemble technique used for multiple classifier ensembling are majority voting, weighted voting, bagging and boosting. The general architecture for cloud classification using two and three ensemble pretrained networks are shown in fig. 2 and fig. 3 respectively.

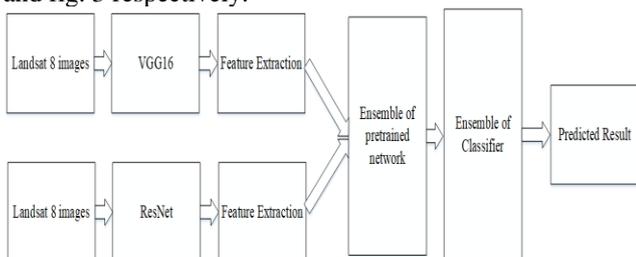


Fig. 2 Ensemble architecture for cloud classification using two pretrained networks

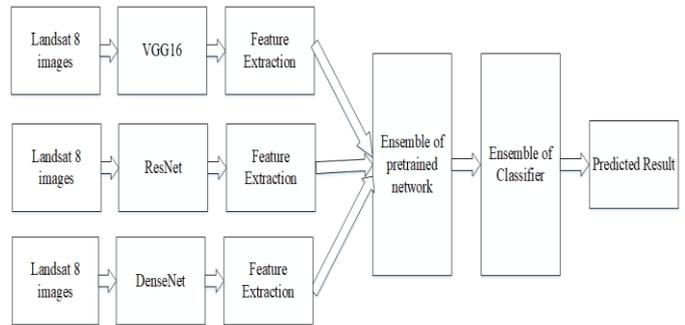


Fig. 3 Ensemble architecture for cloud classification using three pretrained networks

The general steps involved for cloud classification are as follows:

- 1) Landsat 8 image are pre-processed to develop the dataset.
- 2) The dataset is fed to single or ensemble of multiple pretrained neural networks for feature extraction.
- 3) The extracted features are classified using single or ensemble of multiple classifiers.
- 4) Finally, the class of each image is predicted.

2.1 Ensemble of pretrained neural networks:

Ensemble deep learning is the process of combining features of multiple neural networks to improve the performance of the models and reduce the generalization error. The multiple networks extract the diverse features and the features are combined together using classifiers to make the predictions. Ensemble deep learning is a two-step process in which first features are extracted using pretrained networks and these features are predicted using classifiers. The ensemble of the networks yields better results as compared to single network. A pretrained neural network is a neural network trained on large dataset and the features obtained from training is transferable to new target dataset. These networks are good initializers and perform excellent as compared to training the new dataset from scratch with random initial weights. The weights are initialized by training the network on millions of image of ImageNet dataset. Fine tuning the last few layers of the pretrained network will extract the meaningful features associated to target dataset. At the end, the features extracted from pretrained neural networks are concatenated. These features are classified using different classifiers. The different architectures of CNN are as follows:

VGG: Researcher from Oxford Visual Geometry Group developed the simple model for object localization and classification task. It took second place in ILSRCR 2014. The network input size is 224x224 and features extracted are 4096. The convolutional layer has small receptive field of size 3x3 with increase in depth upto 16 or 19 layers. Also, it uses 1x1 convolutional layer to increase the non-linearity of the decision function[13]. ResNet: ResNet is winner of ImageNet competition ILSRCR 2015 resolving the problem of vanishing gradient. It has residual block of 3x3 convolutional layers. The input size required is 224x224 and features extracted by layers are 4096. Another advantage of ResNet is use of global average pooling layers to minimize the overfitting.

The computational complexity is less as compared to AlexNet which is 20 times less deep. VGG is 8 times less deep than ResNet. Good performance of ResNet on image recognition and localization tasks showed that depth is of central importance for many visual recognition tasks[14]. DenseNet: DenseNet is winner of the best paper award in CVPR 2017. It introduces the dense convolutional network where each layer connects another subsequent layer. Thus, it has $L(L+1)/2$ direct connections and learns redundant features as they all share common knowledge. It has several advantages like reducing vanishing gradient problem, strengthening feature propagation, encouraging feature reuse and reducing the number of parameters. The input size accepted by input layers is 224x224 [15].

2.2 Ensemble of classifier:

An Ensemble of classifiers are set of classifiers where individual decisions are merged together in certain way to improve classification result [16]. The ensemble of classifiers gives better accuracy as compared to single classifier. The features from pretrained networks are feed to any machine learning algorithm. Support vector machine, logistic regression, K-nearest neighbors, decision tree, etc. are some of the classifier which can be used to improve the deep learning model accuracy. The main aim of the classifiers is to classify the image based on the learned features. Due to ensemble of classifiers, robust and improved classification results are obtained. There are three basic types of ensemble techniques, namely voting, averaging and weighted averaging.

2.3 Voting based technique:

The voting is a simple method of aggregation of multiple model predictions [16]. There are two different types of voting: hard voting and soft voting. In hard voting, the final predictions are based on majority of votes. This kind of voting model is called majority voting. For the classifiers involving both class and prediction probability for each classifier are called soft voting or weighted voting ensemble. In this method, the average of probability from all classes are calculated and final prediction is based on the highest probability class. Fig. 4 shows the neural network that is used for extracting the features using convolution layer. It extracts the features with respect to edges, contours, texture, etc. The extracted features are forwarded to classifiers for classification.

The final prediction is based on the maximum computed probability of the class.

$$E(X) = \text{arg } g_{i=1,2,\dots,n} \max(C_i(X))$$

Where n represents number of classes and $C_i(X)$ is highest probability class.

The majority and weighted voting final predictions can be computed by following equations:

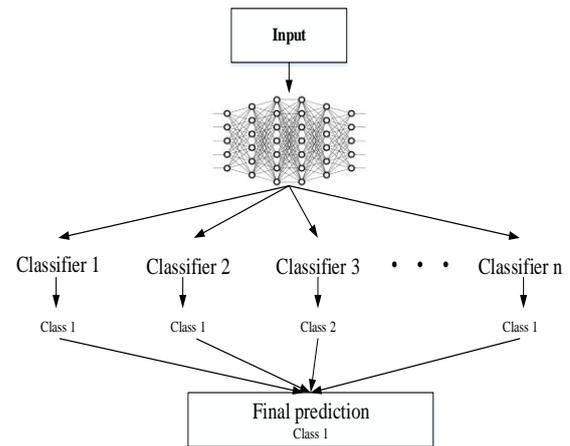


Fig. 4 Ensemble of classifiers using Majority voting technique

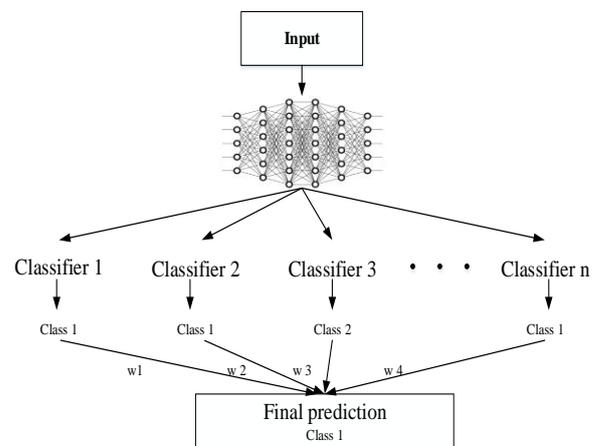


Fig.5 Ensemble of classifiers using Weighted voting technique

$$E(X) = \text{arg } g_{i=1,2,\dots,n} \max \left\{ K_i = \sum_{j=1}^m (C_j(X)) \right\}$$

$$E(X) = \text{arg } g_{i=1,2,\dots,n} \max \left\{ K_i = \sum_{j=1}^m w_j \cdot (C_j(X)) \right\}$$

Where m denotes the number of learners in the classifier ensemble.

2.3.1 Bagging Technique

Bagging is combination of Bootstrap and aggregating ensemble learning method which reduces variance and improves the predictive performance of the classification dataset [18]. In bagging, each classifier selects random data from original dataset to form new dataset which is bootstrapped. Bagging resamples the training set with replacement. This technique works better for unstable learning algorithm.

Algorithm for bagging:

Input: An original training dataset $T = \{(x_1, y_1), (x_2, y_2), \dots, (x_s, y_s)\}$ of size s , W =weak learners, N = number of iterations

Output: Bagged classifier: $G(x) = \text{sign}(\sum_{n=1}^N h_n(x))$

where, $h_n \in [1, -1]$ are the induced classifiers

for $n = 1$ to N do

$T_n \leftarrow \text{Random_sample_Replacement}(s, T)$

$h_n \leftarrow W(T_n)$

end of for loop

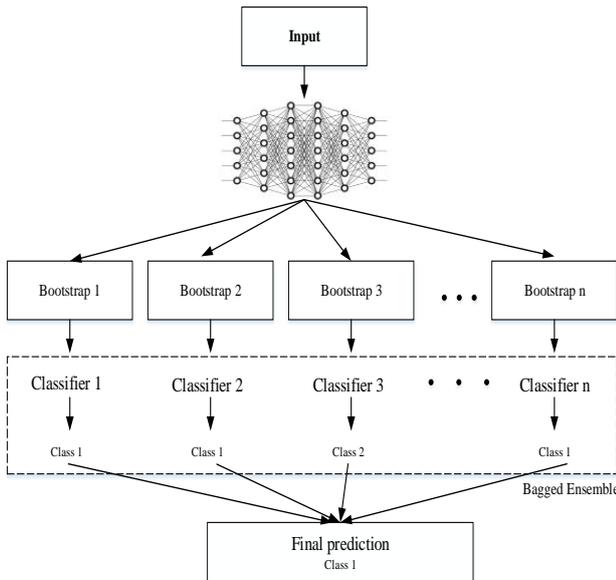


Fig. 6 Ensemble of classifier using Bagging technique

2.3.2 Boosting

Boosting is a meta-algorithm which aims at reducing bias by ensembling the weak learners to form them into strong learners [19]. Boosting is a powerful technique which is similar to bagging technique. The model is trained on the dataset with different classifiers, the poorly predicted results are given more weight and trained again, thus giving final model prediction a high accuracy. Adaboost (adaptive boosting) was first developed by Freund and Schapire in 1996 [20] which is most popular boosting algorithm. It is an iterative ensemble algorithm that improves the accuracy of the classification. Initially, it assigns equal weights to all training dataset. After each round the weights of wrongly classified images are updated so as to increase the accuracy of classification. The boosting helps to efficiently boost the weak learner into a strong learning algorithm by adjusting adaptively the weights of weak learners. The algorithm for boosting is as follows:

Algorithm for boosting:

Input: An original training dataset $T = \{(x_i, y_i)\}$, where $i = 1, 2, 3, \dots, K$, W =weak learners, N = number of iterations

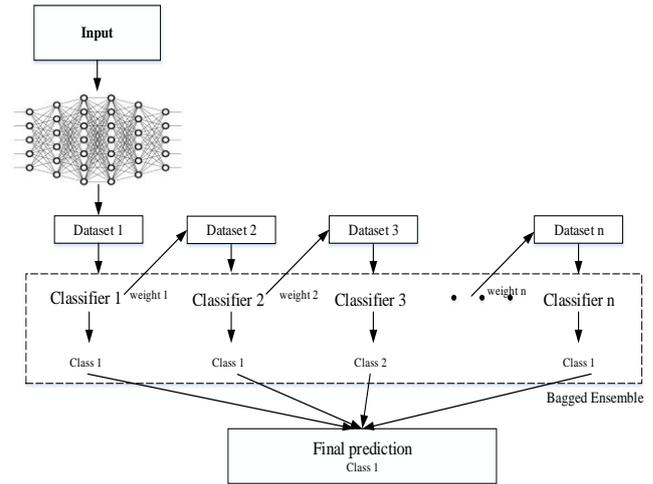


Fig. 7 Ensemble of classifier using Boosting technique

Output: Boosted classifier: $E(x) = \text{sign}(\sum_{n=1}^N \alpha_n h_n(x))$

where, $h_n \in [1, -1]$ are the induced classifiers and α_n are their assigned weights.

- 1: $D_1(i) \leftarrow \frac{1}{K}$ for $i = 1, 2, 3, \dots, K$
- 2: for $n = 1$ to N do
- 3: $h_n \leftarrow W(T, D_n)$
- 4: $\epsilon_n \leftarrow \sum_{i, y_i \neq h_n(x_i)} D_n(i)$
- 5: ϵ_n is weighted error rate of weak classifier,
- 6: if $\epsilon_n > 0.5$ then
- 7: $N \leftarrow n-1$
- 8: return
- 9: end if
- 10: $\alpha_n = \frac{1}{2} \ln\left(\frac{1-\epsilon_n}{\epsilon_n}\right)$
- 11: $D_{n+1}(i) = D_n(i) \cdot e^{-\alpha_n h_n(x_i) y_i}$
- 12: Normalize D_{n+1} to be a proper distribution
- 13: end for loop

III. RESULTS AND DISCUSSION

The proposed cloud classification algorithm was carried out in python Keras library with TensorFlow as backend using GeForce GTX 1080 Ti CUDA enabled Nvidia GPU, 128GB. The dataset of cloud and non-cloud satellite images has 1558 labelled images for training and 589 images for testing. The size of image is 500x500. The primary aim of this paper is to test the proposed algorithms for cloud classification. This can be evaluated by using overall accuracy. Let TP is true positive, TN is true negative, FP is false positive and FN is false negative.

$$\text{Overall accuracy} = \frac{TP + TN}{TP + TN + FN + FP}$$

Table 1 Classification accuracy in percentage

Model	Majority Voting accuracy in %	Weighted Voting accuracy in %	Bagging accuracy in %	Boosting accuracy in %
VGG16	89.56	90.06	92.24	93.65
ResNet	90.65	91.89	93.46	94.90
DenseNet	90.73	92.84	93.29	94.11
Ensemble-1	91.14	94.86	95.98	96.02
Ensemble-2	93.65	96.75	97.73	97.90

First, VGG16 single pretrained network is used as feature extractor and SVM and K-NN are used as classifiers. The classifiers are ensemble together using majority voting. This is giving an accuracy of 89.56%. The weighted voting accuracy is 90.06%. Bagging and boosting have accuracy of 92.24% and 93.65% respectively. Thus, it shows that boosting based technique has good accuracy as compared to other techniques. In same way, for single pretrained networks ResNet and DenseNet with SVM and K-NN classifiers the cloud classification results are also shown in Table 1.

Later, ensemble of VGG16 and ResNet (Ensemble 1) is performed to improve the overall accuracy of classification. The different features extracted from two pretrained networks VGG16 and ResNet are concatenated together. The features are classified using SVM and decision tree classifiers giving better accuracy as compared to single pretrained network. The accuracy has improved up to 96.02%. An experiment with the three pretrained network features extracted from VGG16, ResNet and DenseNet (Ensemble-2) being concatenated and classified using SVM and decision tree is performed. This is yielding best accuracy of 97.90%. Thus, from Table 1, it can be observed that classification accuracy has improved with ensemble of different pretrained networks and ensemble of classifiers. Thus, the ensemble technique with a combination of multiple neural networks and multiple classifiers is able to perform far better in terms of accuracy as compared to single neural network.

IV. CONCLUSION

In this paper the pretrained models VGG16, ResNet, DenseNet were proposed to extract the useful features for a dataset of cloud classification. Ensemble of classifiers such as support vector machine, k-nearest neighbors and decision tree were utilized. While classifying voting, bagging and boosting techniques were employed. The proposed techniques are able to extract diverse features and improve the classification accuracy by reducing the error. The single pretrained network gives an accuracy of 89.56% with VGG16. The accuracy of the majority voting, weighted voting and bagging techniques for ensemble of three pretrained networks are 93.65%, 96.75% and 97.73% respectively. For ensemble using boosting technique, the accuracy has improved to 97.90%. Thus, by combining the pretrained neural networks and classifiers the classification accuracy has improved at the expense of increased complexity.

ACKNOWLEDGMENT

“The authors would like to thank Digital Signal Processing Laboratory, Vellore Institute of Technology, Vellore, India for providing GPU system”

REFERENCES

1. K. G. Crane and R. G. Barry, “The influence of clouds on climate with A focus on high latitude interactions,” *J. Climatol.*, vol. 4, no. 1, pp. 71–93, Jan. 1984.
2. D. P. Y. Suseno and T. J. Yamada, “Two-dimensional, threshold-based cloud type classification using MTSAT data,” *Remote Sens. Lett.*, vol. 3, no. 8, pp. 737–746, Dec. 2012.
3. A. Mefti, A. Adane, and M. Y. Bouroubi, “Satellite approach based on cloud cover classification: Estimation of hourly global solar radiation from meteosat images,” *Energy Convers. Manag.*, vol. 49, no. 4, pp. 652–659, Apr. 2008.
4. Q. Li, W. Lu, and J. Yang, “A Hybrid Thresholding Algorithm for Cloud Detection on Ground-Based Color Images,” *J. Atmos. Ocean. Technol.*, vol. 28, no. 10, pp. 1286–1296, Oct. 2011.
5. T. Inoue, “A cloud type classification with NOAA 7 split-window measurements,” *J. Geophys. Res.*, vol. 92, no. D4, p. 3991, 1987.
6. U. Amato et al., “Statistical cloud detection from SEVIRI multispectral images,” *Remote Sens. Environ.*, vol. 112, no. 3, pp. 750–766, 2008.
7. H. Y. Cheng and C. C. Yu, “Block-based cloud classification with statistical features and distribution of local texture features,” *Atmos. Meas. Tech.*, vol. 8, no. 3, pp. 1173–1182, Mar. 2015.
8. H. Bischof, A. J. Pinz, and W. Schneiden, “Multispectral Classification of Landsat-Images Using Neural Networks,” *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 3, pp. 482–490, 1992.
9. P. D. Heermann and N. Khazenie, “Classification of multispectral remote sensing data using a back-propagation neural network,” *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 1, pp. 81–88, 1992.
10. A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” in *Proc. {NIPS}*, 2012.
11. M. D. Zeiler and R. Fergus, “Visualizing and Understanding Convolutional Networks,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8689 LNCS, no. PART 1, 2014, pp. 818–833.
12. C. Szegedy et al., “Going deeper with convolutions,” in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, vol. 07-12-June, pp. 1–9.
13. K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” pp. 1–14, 2014.
14. 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, vol. 2016-Decem, pp. 770–778.
15. G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, “Densely Connected Convolutional Networks,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, vol. 2017-Janua, pp. 2261–2269.
16. T. G. Dietterich, “Ensemble Methods in Machine Learning,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 1857 LNCS, 2000, pp. 1–15.
17. A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, “A Survey of the Recent Architectures of Deep Convolutional Neural Networks,” pp. 1–67, 2019.
18. L. Breiman, “Bagging predictors: Technical Report No. 421,” *Dep. Stat. Univ. Calif.*, no. 2, p. 19, 1994.
19. R. E. Schapire, “The Boosting Approach to Machine Learning: An Overview,” *Nonlinear Estim. Classif.*, vol. 171, pp. 149–171, 2003.
20. Y. Freund and R. E. Schapire, “A decision-theoretic generalization of on-line learning and an application to boosting,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 1995, vol. 904, no. 1, pp. 23–37.

AUTHORS PROFILE



Ms. Aarti Kumthekar, Received B. E in Electronics and Communication Engineering in 2011 and completed M.E in 2013 from BVCOE, Kolhapur. She is currently doing her research in VIT Vellore. Her research interest Image processing, Satellite Image processing, Machine learning and Deep learning.



G. Ramachandra Reddy, (SM'91) received the M.Sc. and M.Sc. (Tech.) degrees from the Birla Institute of Technology and Science, Pilani, India, in 1973 and 1975, respectively, and the Ph.D. degree from the Indian Institute of Technology Madras, Chennai, India, in 1987. From 1976 to 2010, he was with the Department of Electrical and Electronics Engineering, College of Engineering, Sri Venkateswara University, Tirupati India. From

February 1989 to May 1991, he was Visiting Scientist with the Department of Electrical and Computer Engineering, Concordia University, Montreal, QC, Canada. Currently, he is a senior Professor of School of Electronics Engineering, Vellore Institute of Technology, Vellore, India. Prof. Reddy is a Fellow of the Institution of Electronics and Telecommunication Engineers and a member of the Indian Society for Technical Education.