









# Medical Plants Identification Using Leaves Based on Convolutional Neural Networks

B Ch S N L S Sai Baba<sup>(✉)</sup> , Mudhindi Swathi , Kompella Bhargava Kiran ,  
B. R. Bharathi , Venkata Durgarao Matta , and CH. Lakshmi Veenadhari 

Computer Science and Engineering Department, Vishnu Institute of Technology, Bhimavaram,  
Andhra Pradesh, India  
sai.ossr524@gmail.com

**Abstract.** The ayurvedic medicines have played a crucial role in health system, only a few experts could identify the herbs and know the ayurvedic properties of these herbs. These medicines prepared from the herbs having less side effects as compared to other general medicines. Most of the patients and general medicine users with different diseases are not unaware of the existence of herbal plants and their medical uses and benefits. To make ease of identifying the plants and its medical properties based on the leaf structure, authors developed a system having three architectures which works with Convolutional Neural Networks. Resnet-18, Resnet-50, MobileNet-V2 architectures were used in freeze and unfreeze layers settings. Authors considered ten different kinds of herbal leaves for implementation of the system, in which two thirds of the data used for training and one third for testing. The overall performance of this architecture is checked using accuracy measure and it is observed that three models with freeze layers were showing good performance. Out of these three architectures, Resnet-50 shown accuracy of 95.33%.

**Keywords:** Image Processing · CNN · Medical Plants · Computer Vision

## 1 Introduction

Herbal plants symbolize biodiversity, which is frequently employed as an alternative to conventional medicine. Almost every part of a herbal plant, notably the leaves, can be employed as a component in traditional medicine. This is thus because, in contrast to fruit and roots, leaves are simpler to acquire. Herbal plants play a crucial part in maintaining human health [1]. Given that medical treatment is not accessible to everyone and is expensive in comparison to medical treatment, nearly 80% of people still rely on traditional medicine [2]. For ages, people have employed herbal plants to prevent illness and treat it [3]. Due to the variety of therapeutic plant kinds [4] and the difficulty in differentiating between them, the public is currently unaware of the existence of herbal plants. To recognize and distinguish between these kinds of herbal plants, one needs significant expertise and information. It is important to conserve knowledge of herbal plants so that people can more easily identify the different kinds and use them when

necessary [5]. In general, plant species are still identified manually by comparing and recognizing photos, particularly leaf photographs when the data is already known with the leaves on the plant. Manual identification, meanwhile, still leaves room for mistakes. This is due to the nearly identical leaf color and similar texture and shape of various varieties of herbal plants. Botanists and other individuals with specialized knowledge of herbal plants are the only ones qualified to present different varieties of herbal plants in this manner.

Since not everyone has the requisite knowledge and experience in this area, the community views this method of introducing herbal plants as ineffective for differentiating between the various types of herbal plants [6].

Numerous researchers have used technological advancements and identification of these plants' leaves to conduct research on the classification of herbal plants. Herbal leaf identification is simpler since leaves are more important and accessible than roots [7], which are a portion of the plant that are buried in the ground [8]. Consequently, a system for intelligent and precise herbal leaf identification is required. Most of the earlier research has been devoted to identifying leaves. Numerous classification techniques are employed in the identification system to assist users in recognizing herbal leaves without specialized botanical or anatomical knowledge by identifying the types of herbal plants through leaf identification.

## 2 Related Work

Recently, research has been done to identify the various kinds of herbal leaves. The study by [9] selected therapeutic herbs in order to assess their content value. Computer vision-based feature extraction was used in this investigation. Using ITS2 Sequences and Multiplex-SCAR Markers, [10] authenticated herbal aralia plants carried out another experiment. A categorization of herbal leaves using the SVM method was also done in a different study [11]. Scale Invariant Feature Transform (SIFT) technology was used to extract the picture feature. To combat affine transformations, noise, and changing lighting, the SIFT functionality was deployed. Laws' mask analysis and SVM as the classifier were also used to classify 5 different types of leaves [12]. The obtained accuracy is 90.27%. Using the CNN approach, the identification of Thai medicinal herbs was carried out in [13]. In this paper we are going to train our model on leaves of plants which are mainly found in India. 10 types of herbal plants are considered for this approach. These are Apta: Bauhinia Racemosa, vad: Ficus Benghalensis, Indian Rubber Tree: Ficus Elastica Roxb, Ex Hornem, Karanj: Pongamia Pinnata, Kashid: Senna Siamea (lam.) Irwin & Barneby, Sita Ashok: Saraca Asoka (roxb.) Willd, Pimpal: Ficus Religiosa, Nilgiri: Eucalyptus Globulus, Sonmohar: Peltophorum Pterocarpum, Villayati Chinch: Pithecellobium Dulce [14]. The approach uses Resnet18, Resnet50, MobilnetV2 and MobilenetV3 convolutional neural network to build the models.

## 3 Architectures

In this paper, 3 architectures with 2 variations each are used. They include Resnet 18 with freeze, Resnet 18 without freeze, Resnet 50 with freeze, Resnet 50 without freeze.

### 3.1 Resnet 18

In the article “Deep Residual Learning for Image Recognition” by Kaiming He et al., ResNet-18, a convolutional neural network architecture, was introduced [15]. It is one of the most compact versions of the ResNet model family, which is renowned for its excellence in image recognition tasks.

18 layers make up the ResNet-18 architecture, containing 16 convolutional layers and 2 fully linked layers. The network can learn more effectively thanks to the architecture’s use of residual connections, which propagate information between the layers without causing information loss. In order to do this, skip connections that omit one or more network tiers are added.

The first layers of ResNet-18, such as the convolutional layers, batch normalization, and activation functions, are comparable to those of other convolutional neural networks. ResNet-18’s usage of residual blocks, which are made up of two or more convolutional layers and a skip connection that omits one or more of the layers in the block, is what makes it special.

Multiple image recognition tasks, such as classification, object detection, and segmentation, have been accomplished using ResNet-18. On numerous benchmark datasets, including ImageNet, CIFAR-10, and CIFAR-100, it has been demonstrated to produce state-of-the-art results. It is a well-liked option for applications where computer resources are scarce due to its tiny size and relatively low computational complexity.

### 3.2 Resnet 50

Convolutional neural network ResNet-50 was first presented in the article “Deep Residual Learning for Image Recognition” by Kaiming He et al. [16]. It is one of the more robust models in the ResNet family, which is renowned for its excellence in image recognition tasks.

50 layers make up the ResNet-50 architecture, containing 48 convolutional layers and 2 fully linked layers. It makes use of residual connections, which help the network learn more quickly by transferring knowledge without losing it as it moves through the layers. In order to do this, skip connections that omit one or more network tiers are added.

ResNet-50 is more sophisticated than ResNet-18 and can therefore learn more intricate features from the data. It can also learn more sophisticated representations of the input data because it has larger residual blocks, more convolutional layers, and filters.

Multiple image recognition tasks, such as classification, object detection, and segmentation, have been accomplished using ResNet-50. On numerous benchmark datasets, including ImageNet, COCO, and PASCAL VOC, it has been demonstrated to produce state-of-the-art results. It is a suitable option for activities requiring a high degree of precision and processing resources because of its size and computational complexity.

Overall, ResNet-50 is a potent and popular deep learning model that has shown promise in a variety of image identification applications.

### 3.3 Mobilenet V2

In 2018, Google unveiled MobileNetV2, a convolutional neural network architecture [17]. It is specifically made for embedded and mobile devices, and it is optimized for high precision and minimal computational expense.

By combining depthwise separable convolutions with linear bottlenecks to lower the network's computing cost, MobileNetV2 builds on the original MobileNet architecture. In depthwise separable convolutions, each input channel is subjected to a single filter first, and then the output channels are combined using a 1x1 filter in a pointwise convolution. Comparatively speaking to conventional convolutions, this has fewer parameters and lower processing costs.

To boost the network's nonlinearity without introducing new parameters, linear bottlenecks are used. They are made up of a 1x1 convolution followed by a ReLU activation function, a 3x3 depthwise separable convolution, and a final 1x1 convolution. By reducing the number of parameters and computational cost, the network can learn more intricate features.

To make better use of the network's capacity, MobileNetV2 also has a feature known as inverted residuals. The linear bottleneck, expansion layer, and subsequent linear bottlenecks make up an inverted residual. The expansion layer expands the number of data channels. Due to this, the network can learn more complicated features without having to add more parameters or pay for more compute.

It has been demonstrated that MobileNetV2 can perform at the cutting edge on a range of image recognition tasks, including segmentation, object detection, and classification. It is frequently utilized in embedded and mobile applications when high accuracy is needed but computational resources are constrained.

## 4 Experimental Setup

The leaf dataset consists of 3000 images with 10 classes and each class has 300 images. We have trained our models using 70% of data and reserved 30% for validation. Before applying the model, preprocessing is performed on the dataset. At first, input images are scaled to  $224 \times 224$ , then these images are randomly flipped horizontally with a frequency of 0.5 when using the RandomHorizontalFlip transformation [18]. It indicates that there is a 50% likelihood that each image will be horizontally inverted. The image is horizontally mirrored yet the alteration does not alter the image's content. Each pixel in the image is represented as a floating-point value between 0 and 1, with 0 denoting black and 1 denoting white, and it is converted from a NumPy array or PIL image object into a tensor object. Using mean and standard deviation values, the "Normalize" transformation normalizes the image's pixel values. Based on the statistics of the dataset used for model training, the mean and standard deviation values are computed. The Normalize transformation ensures that the input data have similar statistical features to the data used to train the model by normalizing the input picture. This may aid in enhancing the model's generalizability and accuracy. The dataset has 10 folders with name of the class and has 300 images of that class. By using split function, the data is divided into features and classes. In this model we have initialized the weights from IMAGENET1K\_V1 [19]. A deep learning model's pre-trained weights that have been

learned on the IMAGENET1K\_V1 dataset for image classification tasks are referred to as “IMAGENET1K\_V1 weights.” When a new deep learning model is created and trained on a similar dataset, these weights are often utilized to set its model parameters. The learnt parameters of the deep learning model make up the weights, which are typically in the form of a sizable file. These parameters, which are determined during the training phase on the IMAGENET1K\_V1 dataset, include the weights and biases of the various layers of the model. In order to prevent gradients from earlier iterations from interfering with the gradients calculated for the current batch of data, we first set all gradients to zero. All the network must be frozen here, except for the top layer. In order to prevent the gradients from being computed in backward(), we must set “requires\_grad = False” to freeze the parameters. For the model’s optimization, stochastic gradient descent is applied. The model parameters are updated using tiny batches of data rather than the complete dataset at once, which is a variation on the gradient descent approach. When using SGD, the model parameters are adjusted in the direction of the loss function’s inverse gradient with respect to the parameters.

$$\theta = \theta - \eta \cdot \nabla_{\theta} J(\theta; x^{(i)}; y^{(i)}) \quad (1)$$

where,  $J(\theta)$  is the objective function,  $\eta$  is learning rate,  $\nabla$  is the gradient of objective functions. Especially for big datasets, this speeds up and improves the computational efficiency of computing the gradient. One of the hyperparameters that controls how big of a step is taken in the direction of the negative gradient is learning rate. It is commonly set to a low value, like 0.01 or 0.001, and can be changed during training to enhance the performance of the model. In this paper, the learning rate is set to 0.001[20]. We have used LR scheduler to adjust the learning rates. A learning rate scheduler is a method for modifying the learning rate during training to enhance the model’s performance [21]. A hyperparameter called learning rate regulates how big of a step the optimization method takes when updating parameters [22]. We have used 25 epochs to train the model. The loss of the model is calculated using Cross Entropy Loss [23]. To calculate the cross-entropy loss, the function combines the softmax function and the negative log likelihood loss function. The softmax function turns a set of logits, or unnormalized scores, from the model’s output into a probability distribution over the classes.

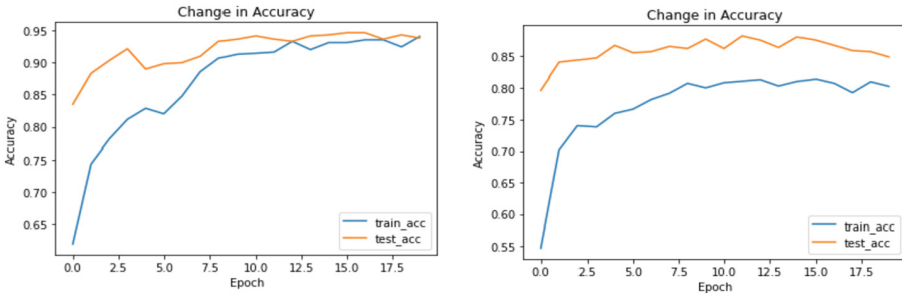
$$f(s)_i = \frac{e^{S_i}}{\sum_j^C e^{S_j}} CE = -\sum_i^C t_i \log(f(s)_i) \quad (2)$$

where,  $S_i$  is input in the form of one hot encoded matrix,  $e$  is the exponent,  $C$  is number of classes,  $t_i$  and  $S_i$  are the ground through variables. The difference between the target class’s actual probability distribution and the anticipated probability distribution is then measured by the negative log likelihood loss function. The model is then connected to a fully connected network which has 10 outputs.

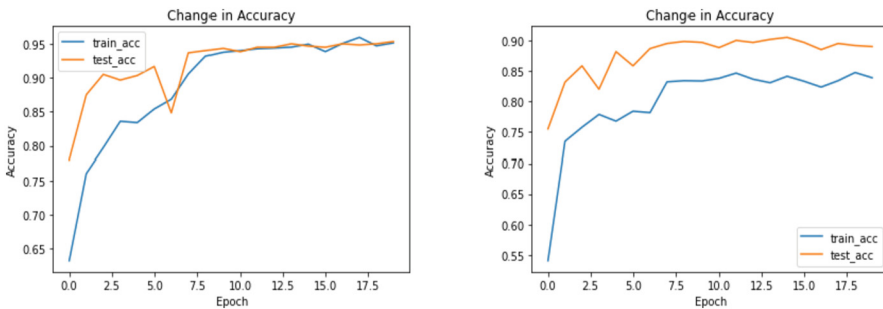
## 5 Results

By using the above setup, Resnet 18 architecture is used to build the model. The accuracy of model stands at 88.16% as in Fig. 1 (left). A variation of the above model is also built as shown in Fig. 1 (Right). In this case, freezing of the layers is not performed to block

backpropagation. In this case the accuracy 94.16%. The same setup is used with Resnet 50 architecture as in Fig. 2. The accuracy of model with and without accuracy is 90.50% and 95.33% respectively. In case of MobileNet V2, we have not used bias and there is no fully connected layer. The accuracy of the model with freeze is 92.83%. In the same way, the accuracy is 93.66% without freezing of layers.



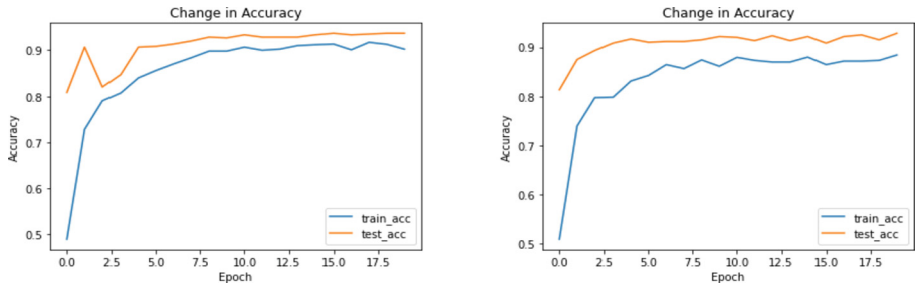
**Fig. 1.** (Left) Accuracy w.r.t to number of epochs using Resnet 18 architecture without freeze. (Right) Accuracy w.r.t to number of epochs using Resnet 18 architecture with freeze.



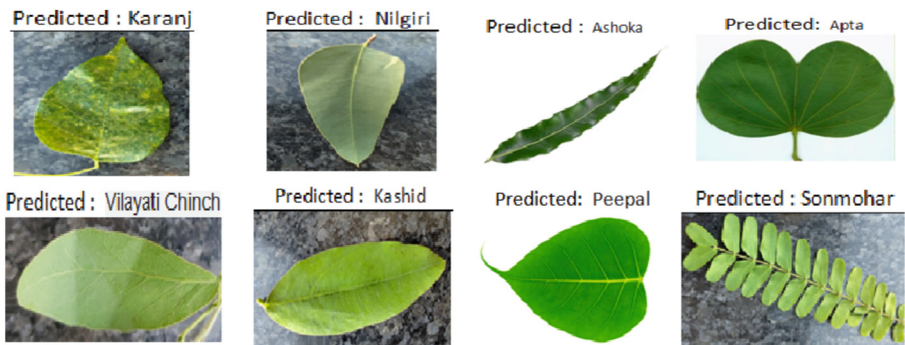
**Fig. 2.** (Left) Accuracy w.r.t to number of epochs using Resnet 50 architecture without freeze. (Right) Accuracy w.r.t to number of epochs using Resnet 50 architecture with freeze.

Figure 3 demonstrates the accuracy of MobileNet V2 w.r.t to number of epochs. The graphs demonstrating the models' accuracy in relation to the number of epochs are shown above. We have always considered 20 epochs because overfitting has been shown above this. Resnet 50 without freezing has outperformed among all the models. The model's output, which is displayed below in Fig. 4, uses plant identification based on leaves.

The performance of all models, both with and without freezing layers, is summarized in the table below with respect to the number of parameters, epochs, the optimizer, and accuracy (Table 1).



**Fig. 3.** (Left) Accuracy w.r.t to number of epochs using MobileNet-V2 architecture without freeze. (Right) Accuracy w.r.t to number of epochs using MobileNet-V2 architecture with freeze.



**Fig. 4.** Shows sample predictions made by the above architectures.

**Table 1.** Model performance w.r.t to number of parameters

Architectures	Total Parameters	Trainable Parameters	Epochs	Optimizer	Accuracy
Resnet-18 with freeze	11,181,642	0	20	SGD	88.16
Resnet-18 without freeze	11,181,642	5,130	20	SGD	94.16
Resnet-50 with freeze	23,528,522	20,490	20	SGD	90.50
Resnet-50 without freeze	23,528,522	23,528,522	20	SGD	95.33
MobileNet-V2 With freeze	2,236,682	1,694,154	20	SGD	92.83
MobileNet-V2 without freeze	2,236,682	2,236,682	20	SGD	93.66

## 6 Conclusion and Future Scope

We were able to identify 10 different kinds of leaves of plants having good medical properties. Most of the plants are used in Ayurveda for curing many diseases. Based on 3000 photos of medicinal plants, we developed 6 image classification models (Resnet-18 with freeze, Resnet-18 without freeze, Resnet-50 with freeze, Resnet-50 without freeze,

MobileNet-V2 With freeze, MobileNet-V2 without freeze) for this study. The dataset of leave images is publicly available. We found that Resnet- 50 architecture without freeze layers has outperformed other models. We have used accuracy to measure the performance of the model and got 95.33% in this setting. By creating a mobile application and integrating more classes of leaves, this study can be furthered.

## References

1. Sofowora, A., Ogunbodede, E., Onayade, A.: The role and place of medicinal plants in the strategies for disease prevention. *Afr. J. Tradit. Complement. Altern. Med.* **10**(5), 210–229 (2013). <https://doi.org/10.4314/ajtcam.v10i5.2>
2. Ekor, M.: The growing use of herbal medicines: issues relating to adverse reactions and challenges in monitoring safety. *Front. Pharmacol.* **4**, 177 (2014). <https://doi.org/10.3389/fphar.2013.00177>
3. Eddouks, M., Chattopadhyay, D., De Feo, V., Cho, W.C.: Medicinal plants in the prevention and treatment of chronic diseases. *Evid Based Complement Alternat Med.* **2012**, 458274 (2012). <https://doi.org/10.1155/2012/458274>
4. Firenzuoli, F., Gori, L.: Herbal medicine today: clinical and research issues. *Evid Based Complement Alternat Med.* **4**(Suppl 1), 37–40 (2007). <https://doi.org/10.1093/ecam/nem096>
5. Chen, S.L., Yu, H., Luo, H.M., Wu, Q., Li, C.F., Steinmetz, A.: Conservation and sustainable use of medicinal plants: problems, progress, and prospects. *Chin Med.* **11**, 37 (2016). <https://doi.org/10.1186/s13020-016-0108-7>
6. Welz, A.N., Emberger-Klein, A., Menrad, K.: Why people use herbal medicine: insights from a focus-group study in Germany. *BMC Complement. Altern. Med.* **18**, 92 (2018). <https://doi.org/10.1186/s12906-018-2160-6>
7. Wachtel-Galor, S., Benzie, I.F.F.: Herbal medicine: an introduction to its history, usage, regulation, current trends, and research needs. In: Benzie, I.F.F., Wachtel-Galor, S. (eds.) *Herbal Medicine: Biomolecular and Clinical Aspects*, Chapter 1. CRC Press/Taylor & Francis (2011). <https://www.ncbi.nlm.nih.gov/books/NBK92773/>
8. Bhat, S.G.: Medicinal plants and its pharmacological values. In: *Natural Medicinal Plants*, IntechOpen, May 11, 2022. <https://doi.org/10.5772/intechopen.99848>
9. Venkataraman, D., Nehru, M.: Computer vision based feature extraction of leaves for identification of medicinal values of plants. In: *2016 International Conference on Computational Intelligence and Communication Networks (CICN)*, pp. 1–5 (2016). <https://doi.org/10.1109/ICCIC.2016.7919637>
10. Kim, W., Moon, B., Yang, S., Han, K., Choi, G., Lee, A.Y.: Rapid authentication of the herbal medicine plant species *aralia continentalis* Kitag. and *Angelica biserrata* C.Q. Yuan and R.H. Shan using ITS2 sequences and multiplex-SCAR markers. *Molecules* **21**, 270 (2016). <https://doi.org/10.3390/molecules21030270>
11. Kaur, P., Singh, S.: Classification of Herbal Plant and Comparative Analysis of SVM and KNN Classifier Models on the Leaf Features Using Machine Learning. (2021). [https://doi.org/10.1007/978-981-16-1048-6\\_17](https://doi.org/10.1007/978-981-16-1048-6_17)
12. Suh, H., Hofstee, J.W., IJsselmuiden, J., Van Henten, E.J.: Sugar beet and volunteer potato classification using Bag-of-Visual-Words model, Scale-Invariant Feature Transform, or Speeded Up Robust Feature descriptors and crop row information. *Biosyst. Eng.* **166**, 210–226 (2017). <https://doi.org/10.1016/j.biosystemseng.2017.11.015>
13. Mookdarsanit, L., Mookdarsanit, P.: Thai Herb Identification with Medicinal Properties Using Convolutional Neural Network (2019)

14. Rokde, V., Raut, P.: Leaf Images Dataset-Indian Trees leaf's Dataset for image classification problem statement, Version 1 (March 2023). Retrieved: 03/04/2023 from <https://www.kaggle.com/datasets/ichhadhari/leaf-images>
15. Ou, X., et al.: Moving object detection method via ResNet-18 with encoder–decoder structure in complex scenes. *IEEE Access* **7**, 108152–108160 (2019). <https://doi.org/10.1109/ACCESS.2019.2931922>
16. Rifa'i, A.M., Utami, E., Ariatanto, D.: Analysis for diagnosis of pneumonia symptoms using chest X-Ray based on Resnet-50 models with different epoch. In: 2022 6th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), pp. 471–476 (2022). <https://doi.org/10.1109/ICITISEE57756.2022.10057805>
17. Sun, Q., Luo, X.: A new image recognition combining transfer learning algorithm and MobileNet V2 model for palm vein recognition. In: 2022 4th International Conference on Frontiers Technology of Information and Computer (ICFTIC), pp. 559–564 (2022). <https://doi.org/10.1109/ICFTIC57696.2022.10075212>
18. Shorten, C., Khoshgoftaar, T.M.: A survey on image data augmentation for deep learning. *J Big Data* **6**, 60 (2019). <https://doi.org/10.1186/s40537-019-0197-0>
19. Jayakody, D.: Custom Image Classifier with PyTorch – A Step-by-Step Guide. Artificial Intelligence, Computer Vision- URL: <https://dilithjay.com/blog/custom-image-classifier-with-pytorch-a-step-by-step-guide/>
20. Konar, J., Khandelwal, P., Tripathi, R.: Comparison of various learning rate scheduling techniques on convolutional neural network. In: 2020 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), pp. 1–5 (2020). <https://doi.org/10.1109/SCEECS48394.2020.94>
21. Nedunuri, S.U.D.: Crop disease prediction with convolution neural network (CNN) augmented with cellular. *Int. Arab J. Inf. Technol. (IAJIT)* **19**, 69–77 (2022). <https://doi.org/10.34028/iajit/19/5/8>
22. Murthy, M.Y.B., Koteswararao, A., Babu, M.S.: Adaptive fuzzy deformable fusion and optimized CNN with ensemble classification for automated brain tumor diagnosis. *Biomed. Eng. Lett.* **12**, 37–58 (2022). <https://doi.org/10.1007/s13534-021-00209-5>
23. Rezaei-Dastjerdehei, M.R., Mijani, A., Fatemizadeh, E.: Addressing imbalance in multi-label classification using weighted cross entropy loss function. In: 2020 27th National and 5th International Iranian Conference on Biomedical Engineering (ICBME), pp. 333–338 (2020). <https://doi.org/10.1109/ICBME51989.2020.9319440>