An efficient multi-functional deep learning model for effective medical image classification using skin lesion database

Kishore Babu Nampalle
Computer Science and Engineering Department
Machine Intelligence Lab
Indian Institute of Technology Roorkee, India
kbabu89@cs.iitr.ac.in

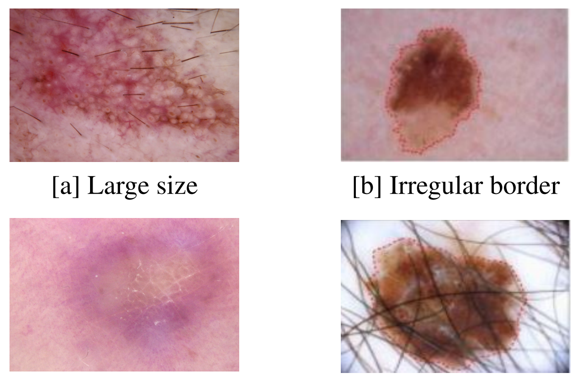
Balasubramanian Raman
Computer Science and Engineering Department
Machine Intelligence Lab
Indian Institute of Technology Roorkee, India
bala@cs.iitr.ac.in

Abstract

The automatic process of classifying a medical image plays a vital role in Computer-Aided Diagnosis (CADx). Due to the advent of Convolutional Neural Networks (CNNs) and wide usage, there has been a substantial improvement in the performance of the classification process combined with the process of implicit feature extraction. CNN requires a large amount of data, but building an extensive data set is challenging. Hence, Transfer learning appeared to resolve the same issue. Predefined models like MobileNet, VGG19, Inception-V3, and ResNet50, based on datasets with more sizes such as ImageNet, play a vital role in training and improving the performance. Extracting such unique features from medical images is a challenging task due to the different properties of images. Training a Deep Neural Network is an intensive task because it requires high configured computing machines and may require more time. Hence, this paper proposed a multi-functional deep learning architecture, including an ensemble of Logistic Regression classifiers and a MobileNet pre-trained model. Here, the input data of skin lesion images from the ISIC challenge dataset for binary and multi-class classification. Obtained results are compared with other models with the help of performance metrics.

1. Introduction

Medical imaging [1] refers to a range of techniques for visualising the human body in order to diagnose, monitor, or treat medical conditions. Extensive knowledge is required to efficiently and correctly interpret the images generated by any of these modalities, which include radiography [4], X-ray imaging [21], Computed Tomography (CT scan) imaging, ultrasound [24], and magnetic resonance imaging. It has been playing a very vital role in the biomedical image processing research domain [11]. It is used to recreate the medical images for clinical diagnosis and further treatment purpose. With the increase in the number of medical imaging modalities [23], there will be a necessity for system setup and experts for clinical diagnosis due to various properties of images as shown in Fig. [1].



[a] Large size [b] Irregular border [c] Low contrast surrounding [d] Other artifacts

Figure 1. Samples of skin lesion images with varying features.

Computer Vision, Pattern Recognition, Picture Mining, and Machine Learning are now all part of 3D medical image processing in various dimensions. Deep learning [3] is a prominent method of determining the validity of a condition, and its applications in healthcare range from cancer detection to infection monitoring to personalized medication recommendations. Clinicians now have access to a wealth of information from various sources, including radiological imaging, genetic sequencing, and pathological imaging.

Transfer Learning [20] is one of the most demanding approaches in computer vision and data science, and it's also essential for feature extraction and learning. Deep Neural

2770-4319/22/$31.00 ©2022 IEEE
DOI 10.1109/MIPR54900.2022.00083

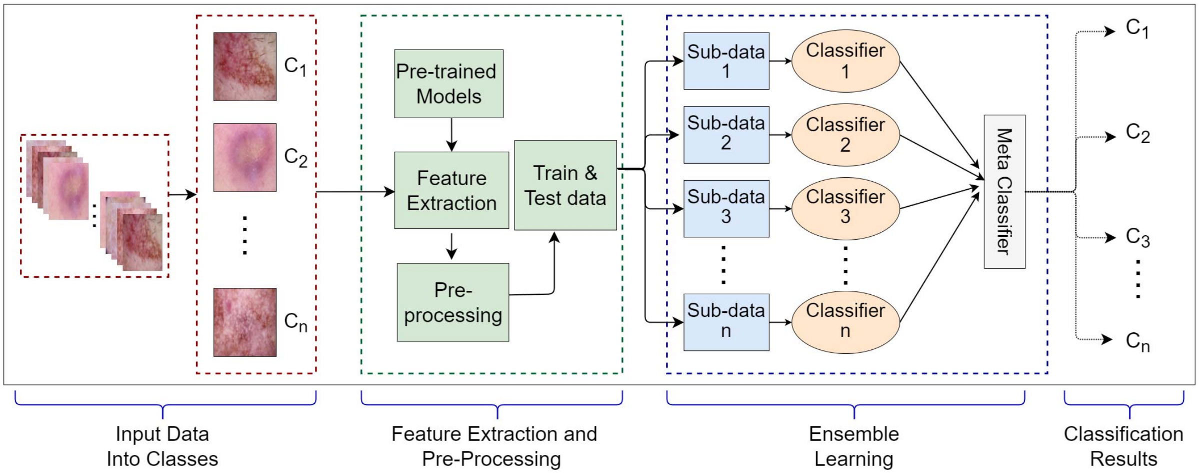


Figure 2. Schematic architecture for the proposed methodology.

Networks (DNNs) have shown to be effective in object processing tasks such as image segmentation, image classification, and feature extraction by requiring vast amounts of data to train the model [16, 12].

Semi-supervised learning approaches have been developed in medical image analysis for a long time to reduce human effort [5, 19]. The training method, which iteratively updates network parameters and labels for unlabeled data, has been used in a few medical image processing applications [8][6]. Similarly, for the segmentation of the liver [17] and breast cancer [13], the co-training technique has been implemented, in which various classifiers have been considered and trained on feature sets separately. Due to inter and intra-class [18] characteristics and inter-class variations, it is a more complex task. The outcomes are determined by the likelihood of each classifier's prediction.

The main objective is to prevent overfitting a new problem with small data by using the knowledge of a model that has learned from large-scale labeled data, such as ImageNet. It reduces training time because no more data is required for the training process, and the training process does not have to be restarted from the beginning. The results were analyzed using deep architectures such as MobileNet, ResNet50, VGG19, Inception-ResNet-V2, Xception, and Inception-V3.

The contributions of the paper are summarized as follows.

- The proposed framework of Deep Ensemble Learning uses a supervised learning technique for binary and multi-class classification of skin lesion images from the ISIC dataset.

- Feature extraction has been done using CNN, and pre-processing techniques such as augmentation, data normalization, and morphological operations have been applied to the given dataset.

- Experimental results have demonstrated the effective performance of the proposed method and compared it with other models.

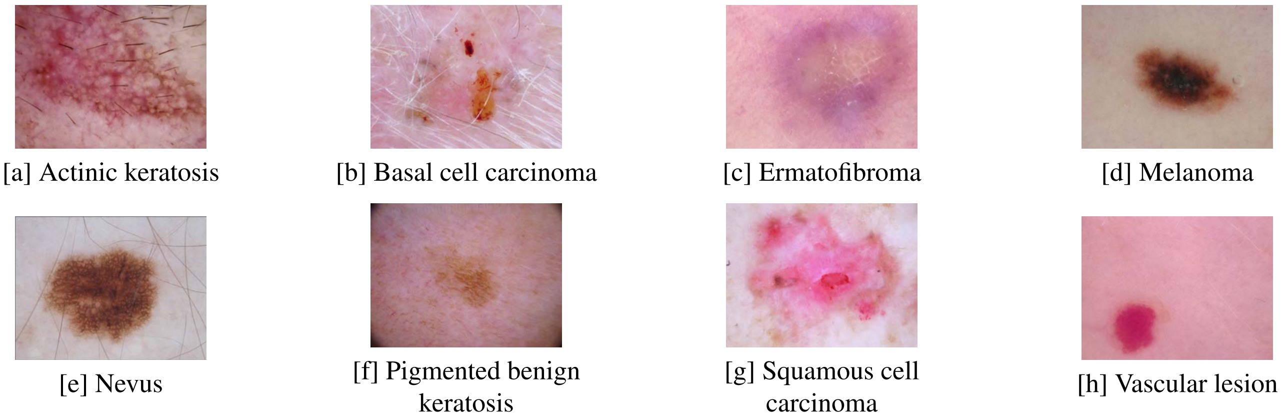
2 PROPOSED METHOD

The proposed method for the classification process is divided into two steps. The first one is feature extraction, and the second one is classifying the image data set using a deep ensemble learning model. The proposed method's architecture is shown in Fig. [2].

2.1 Feature Extraction and Pre-Processing Phase

Several convolution layers are used in feature extraction, followed by max-pooling and an activation function. Convolution layers are described with digital filters and used for image transformation and the pooling layer is used for dimensionality reduction with the help of pixel transformation. Data augmentation is a regularization technique used to generate more training data using a given dataset to reduce overfitting.

Data normalization and other transformations include rotation, changing the contrast, shifting, adjustment of exposure to make it large the size of the dataset, and morphological operations to locate the lesion properly. The input image is given for feature extraction and pre-processing and features have been extracted with the help of MobileNet pre-trained model and given to the next stage of processing.



[a] Actinic keratosis [b] Basal cell carcinoma [c] Ermatofibroma [d] Melanoma

[e] Nevus [f] Pigmented benign keratosis [g] Squamous cell carcinoma [h] Vascular lesion

Figure 3. Samples of dataset for Multi Class Classification.

2.2 Ensemble Learning Phase

Using [22] a single classifier, among the various methodologies for image classification, might sometimes fail to accomplish the task correctly when all features of the data are met. As a result, to overcome these issues, a combination of classifiers is used. Ensemble learning [2] is a meta approach to machine learning that combines predictions from numerous distinct base models to produce predictive performance. It separates the provided dataset (size N) into subsets (size n), where $n < N$, and assigns each base model a subset (size n).

The proposed model's major aim is to reduce overfitting and construct a more flexible model for enhancing classification task performance and results with good predictive accuracy. The classifier's accuracy is determined by its bias and variance values. The discrepancy in prediction between the true model (model constructed on supplied data) and the average model (average of all predictions on various samples of given data) is known as bias . The difference in prediction between all models and the average model is known as variance.

Linear classifiers are less versatile because they have a strong bias. Non-linear classifiers have a low bias, which makes them more flexible, but they have a high variance, which leads to overfitting. As a result, an ensemble learning architecture is proposed for achieving low bias and low variance by combining [2] the base models for improved performance. In this case, the base models are logistic regression classifiers, and the proposed model was implemented using the MobileNet pre-trained model. The performance of ensemble learning over the individual basis model is shown in Eq. [1].

$$\text{A} = 1 - (1 - x)^y \tag{1}$$

where x is the accuracy of each base model, y is the number of base models, and A is the accuracy of ensemble learning.

3 EXPERIMENTS AND RESULTS

The experiments have been performed for binary and multi-class classification using the proposed system.

3.1 Dataset and Training Strategy

The effective performance of the proposed model in the classification of medical images is illustrated with the help of skin lesion images from the ISIC challenge data set and Logistic Regression classifiers. Fig. [4] and Fig. [3] show some of the sample results for binary and multi-class classification.

[a] Malignant [b] Benign

Figure 4. Samples of dataset for Binary Classification.

The data set for multi-class classification has 1600 skin lesion images with 224×224 size and is divided into eight classes equally; each class has 200 images. Another dataset for binary classification has 400 images, among those 200 images are of type Malignant and 200 are of type Bening. Automatic feature learning and pre-processing have been done with the help of a pre-trained model (Mobile Net). The train-test split of 80-20 has been used.

3.2 Metrics and Evaluation

The following metrics have been used to evaluate the proposed system. where 'TP', 'FP', 'TN', and 'FN' denote true positive, false positive, true negative, and false negative respectively. Metrics such as Precision (Eq. 2), Recall (Eq. 3), F1-Score (Eq. 4), Accuracy (Eq. 5) have been used to analyse the performance of the proposed method.

$$\textbf{Precision} = \frac{TP}{TP + FP} \tag{2}$$

$$\textbf{Recall} = \frac{TP}{TP + FN} \tag{3}$$

$$\textbf{F1-Score} = 2\times\frac{P * R}{P + R} \tag{4}$$

$$\textbf{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{5}$$

3.3 Result Analysis

Produced results (average values) using the proposed method are shown below with the help of confusion matrix [5], and performance metrics as shown in tables [1] and [2].The proposed method resulted in an accuracy of 91.50% for binary classification and 86.10% for multi-class classification.

Table 1. Comparison of results for Binary classification.

Method	Precision	Recall	F1-Score	Accuracy
Xception [15]	0.85	0.85	0.85	0.987
Inception-ResNet-V2 [9]	0.86	0.86	0.86	0.851
Inception-V3 [10]	0.89	0.84	0.84	0.882
ResNet50 [7]	0.76	0.76	0.76	0.803
Proposed {MobileNet}	**0.88**	**0.88**	**0.88**	**0.915**

Table 2. Comparison of results for multi-class classification.

Model	Precision	Recall	F1-Score	Accuracy
Inception-V3 [10]	0.83	0.78	0.80	0.82
VGG19 [14]	0.83	0.80	0.81	0.80
ResNet50 [7]	0.80	0.69	0.74	0.79
Proposed {MobileNet}	**0.83**	**0.82**	**0.84**	**0.86**

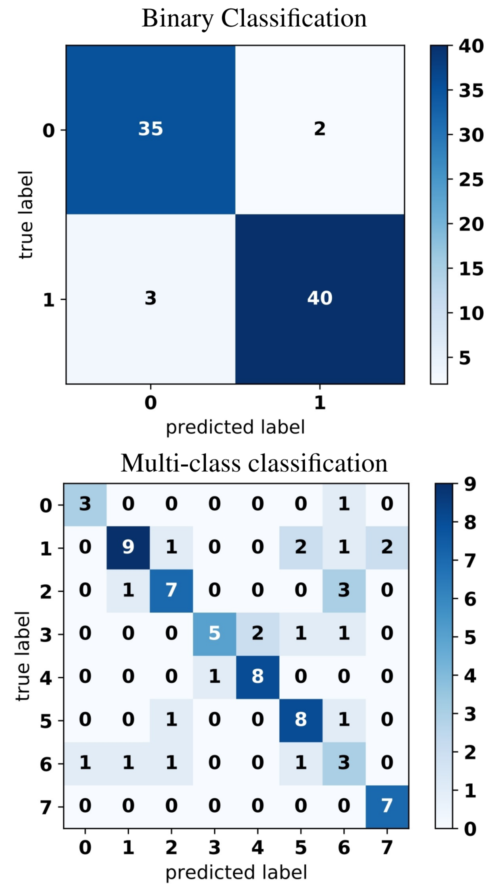


Figure 5. Confusion Matrices for binary and multi-class classification using MobileNet model and Logistic Regression classifier

4 DISCUSSION

Classification is a probabilistic procedure that involves converting a class label and utilising activation functions to predict class membership. The number of classes determines the type of activation function that can be employed. The sigmoid function can be used for binary classification with two mutually exclusive classes. The Softmax can be used for multi-class classification with more than two mutually exclusive classes. The proposed model is used in this paper to classify both binary and multi-class data using a specified activation function. The sigmoid function was utilised for binary classification, and the softmax function was employed for multi-class classification. The proposed model's results (accuracy for binary and multi-class classification) are compared to those of other classifiers such as the Random forest classifier (90%, 84%), Decision trees (85%, 78%), and KNN (86% , 80%), and SVM (89%, 83%).

5 CONCLUSION

This paper proposes an efficient framework of supervised homogeneous ensemble learning that includes an experiment on two datasets (ISIC challenging dataset) of dermoscopic images for binary and multi-class classification. The proposed model produces effective results with performance gain in terms of performance metrics. This approach is more helpful for researchers and clinicians to analyze the skin lesion and further skin diagnosis. Other optimization and adaptive algorithms can be considered in the future, and the latest dataset can be used to estimate the performance. We are planning to extend this work to segmentation tasks.

Acknowledgments

This research was carried out at the Indian Institute of Technology Roorkee's Machine Intelligence Lab.

References

[1] M. Aljabri, M. AlAmir, M. AlGhamdi, M. Abdel-Mottaleb, and F. Collado-Mesa. Towards a better understanding of annotation tools for medical imaging: a survey. *Multimedia Tools and Applications*, pages 1–35, 2022.

[2] W. H. Beluch, T. Genewein, A. Nürnberger, and J. M. Köhler. The power of ensembles for active learning in image classification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 9368–9377, 2018.

[3] X. Chen, X. Wang, K. Zhang, K.-M. Fung, T. C. Thai, K. Moore, R. S. Mannel, H. Liu, B. Zheng, and Y. Qiu. Recent advances and clinical applications of deep learning in medical image analysis. *Medical Image Analysis*, page 102444, 2022.

[4] A. A. Drosos, E. Pelechas, and P. V. Voulgari. A patient with symmetrical polyarthritis. the value of conventional radiography for a correct diagnosis. *Rheumatology and Therapy*, 9(2):771–779, 2022.

[5] B. et al. Semi-Supervised Learning for Network-Based Cardiac MR Image Segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 253–260. Springer, 2017.

[6] B. et al. Self-Supervised Learning for Cardiac (MR) Image Segmentation by Anatomical Position Prediction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 541–549. Springer, 2019.

[7] C. et al. Deep Residual Learning for Image Compression. In *CVPR Workshops*, page 0, 2019.

[8] G. et al. Semi-Supervised Learning for Biomedical Image Segmentation via Forest Oriented Super Pixels (voxels). In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 702–710. Springer, 2017.

[9] K. et al. Automated Diabetic Macular Edema (DME) Analysis using Fine Tuning with Inception-Resnet-V2 on OCT Images. In *2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, pages 442–446, 2018.

[10] M. et al. Transfer Representation Learning using Inception-V3 for the Detection of Masses in Mammography. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2587–2590, 2018.

[11] P. et al. Unsupervised Domain Adaptation for Medical Imaging Segmentation with Self-Ensembling. *NeuroImage*, 194:1–11, 2019.

[12] R. et al. Chexnet: Radiologist-Level Pneumonia Detection on Chest x-rays with Deep Learning. *arXiv preprint arXiv:1711.05225*, 2017.

[13] S. et al. Computerized Breast Cancer Analysis System using Three Stage Semi-Supervised Learning Method. *Computer methods and programs in biomedicine*, 135:77–88, 2016.

[14] S. et al. Multi-Modal Deep Guided Filtering for Comprehensible Medical Image Processing. *IEEE Transactions on Medical Imaging (TMI)*, 39(5):1703–1711, 2020.

[15] T. et al. Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning? *IEEE Transactions on Medical Imaging (TMI)*, 35(5):1299–1312, 2016.

[16] T. et al. The Ham10000 Dataset, A Large Collection of Multi-Source Dermatoscopic Images Of Common Pigmented Skin Lesions. *Scientific data*, 5(1):1–9, 2018.

[17] X. et al. 3d Semi-Supervised Learning with Uncertainty-Aware Multi-View Co-Training. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 3646–3655, 2020.

[18] Y. et al. Automated Melanoma Recognition in Dermoscopy Images via Very Deep Residual Networks. *IEEE transactions on medical imaging (TMI)*, 36(4):994–1004, 2016.

[19] Z. et al. Deep Adversarial Networks for Biomedical Image Segmentation Utilizing Unannotated Images. In *International conference on medical image computing and computer-assisted intervention (MICCAI)*, pages 408–416. Springer, 2017.

[20] Z. et al. A Comprehensive Survey on Transfer Learning. *Proceedings of the IEEE*, 109(1):43–76, 2021.

[21] C. Ieracitano, N. Mammone, M. Versaci, G. Varone, A.-R. Ali, A. Armentano, G. Calabrese, A. Ferrarelli, L. Turano, C. Tebala, et al. A fuzzy-enhanced deep learning approach for early detection of covid-19 pneumonia from portable chest x-ray images. *Neurocomputing*, 2022.

[22] S. Kiruthika, V. Masilamani, and P. Joshi. Fusion of image quality assessment and transfer learning for covid19 detection using ct scan image. In *12th Indian Conference on Computer Vision, Graphics and Image Processing, ICVGIP 2021*, 2021.

[23] P. Kumar, S. Srivastava, and R. Srivastava. Basic understanding of medical imaging modalities. In *High-Performance Medical Image Processing*, pages 1–17. Apple Academic Press, 2022.

[24] J.-Y. Lu, P.-Y. Lee, and C.-C. Huang. Improving image quality for single-angle plane wave ultrasound imaging with convolutional neural network beamformer. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 69(4):1326–1336, 2022.