2021 Sixth International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET) | 978-1-6654-4086-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/WISPNET51692.2021.9419466

Enhanced Dermatoscopic Skin Lesion Classification using Machine Learning Techniques

P. Varalakshmi[1], V. Aruna Devi[2], M. Ezhilarasi[3], and N. Sandhiya[4]

[1,2,3,4]Department of Computer Technology, Madras Institute of Technology, Chennai, India
[1]varanip@gmail.com, [2]arunadev2000@gmail.com,
[3]ezhilarasi3098@gmail.com, [4]sandhiyaguna96@gmail.com

***Abstract*—Malicious melanoma will create greatest impact to the people. Most of the research was done on this field to find out the benign and malignant images and also to classify the different kind of skin lesions. So in the process of dermatoscopic skin lesion classification, segmentation, feature extraction, classification of the dermatoscopic images will play an important role. In this paper, we are focusing on the classification part. We are using the MNIST HAM 10000 dataset. This paper consists of two parts. In the first step, we have analyzed that always the balanced dataset will provide better accuracy compared to the imbalanced dataset. So in order to make the dataset balanced, we have used the up sampling method called Synthetic Minority Oversampling Technique (SMOTE) which greatly improves the accuracy of most of the machine learning models. Then we have analyzed the accuracy of different machine learning algorithms we have implemented. As a result, we have concluded that Support Vector Machine algorithm with Polynomial kernel provides better accuracy compared to other machine learning algorithms like Decision Tree using Gini index and Entropy, Naïve Bayes, XGBoost, Random Forest, Support Vector Machine and Logistic Regression algorithms. We have chosen Precision, Recall, F1-Score as evaluation metrics along with Accuracy. In our system, as the highest we got 96.825% accuracy with Support Vector Machine using Polynomial Kernel. Since XGBoost is a Gradient Boosting algorithm, the result provided by the algorithm might vary. In order to validate the accuracy obtained from the XGBoost algorithm, we have used the k-fold cross validation method in the XGBoost method. In our system, we have used k=10 in k-fold cross validation algorithm and we got an accuracy of 95.984% and we have concluded that SVM with Polynomial kernel provides better accuracy compared to all other algorithms.**

***Index Terms*—Melanoma, Skin Lesion Segmentation, Dermatoscopic Images, Machine Learning, SMOTE, SVM.**

I. Introduction

In the field of medicine, malicious skin lesions classification plays an important role. Typically, skin lesions are different from the actual surrounding area of the skin. Some of the skin lesions are harmless whereas some might be very harmful. Since, finding the skin lesion itself a difficult task, to classify what kind of skin lesion has occurred in the skin will be a challenging task. Owing to the surrounding environment, the noise, hair and color of the person, the task will be much more difficult. The classification of skin lesions include so many steps beginning with imaging and ending with classification. In this paper, we have taken the benchmark dataset MNIST HAM 10000 dataset which consists of 10,015 images which belongs to different skin lesion diseases. In this dataset, there are totally 7 types of skin lesions. For classification, there are so many Machine Learning and Deep Learning algorithms used. Since we have taken the images in the form of pixels in the CSV (Comma Separated Value) file, we have done a comparative analysis by using various Machine Learning Algorithms. In [12], various machine learning algorithms like Multiple Instance Learning techniques, K-Nearest Neighbor Algorithm, Decision Trees, Logistic Regression, Support Vector Machine, Artificial Neural Network and Deep Learning have been used for Automated Melanoma Detection. In [13], Melanoma prediction was done using Deep Learning Techniques, in which the authors have created the model using built-in network VGG-16. We have used the algorithms like Support Vector Machine, Naïve Bayes, Random Forest, Decision Tree using Gini index and Entropy, Logistic Regression, and XGBoost.

In our proposed system, we have done with two things. The first one is about how to improve the accuracy of the skin lesion classification model. For this, we have done the up sampling and it actually provides better accuracy. The second thing was focused on which algorithm provides better results, for this we have compared all the machine learning algorithms and found out that SVM (kernel=Polynomial) provides better accuracy. We have also evaluated the model using Baseline CNN (Convolutional Neural Network) but it provides 82.424% Accuracy only. Since our dataset is in the form of CSV file where images were stored in the form pixel values, we have used Machine Learning models for our system.

The paper is organized into 5 sections. Section I deals with the introduction, section II deals with the literature survey about skin lesion classification. In section III we have proposed our system for skin lesion classification. In section IV, the details about the implementation and results were presented. In section V, the proposed system was evaluated by comparing the accuracy of different machine learning models.

II. Related Works

The interrelationship between the skin lesions and the corresponding informative context was modeled by the bi-directional dermatoscopic feature learning (biDFL) [1]. To obtain informative decisions from multiple classification layers, a multi-scale consistent decision fusion (mCDF) was used by the system proposed in [1]. In order to enhance abstract-level explanative performance the biDFL has acquired the position above all the CNN network layers. This high-level performance was achieved by the particulars of attributes flow via the two supporting directions. For performance evaluation of skin lesion classification they have used G-mean, Dice coefficient, Jaccard index, and segmentation accuracy and as a result, for ISBI 2016 and 2017 databases got 57.8% and 70.3% accuracy.

To make the result of feature extraction better, a CAD system was proposed in [3], in which they have proposed the comparative study on both ABCD and three point checklist methods with the help of WEKA software by using various algorithms like J48 Decision Tree, Random Forest, Sequential Minimal Optimisation Logistic Regression, and they have concluded that they got better accuracy with three point checklist method. But they didn't apply their proposed model with any standard algorithm, they have used the PH2 dataset which has only 200 images with 3 different classes of skin lesions.

In [4], to classify and detect the skin lesions in teledermatology, a system was proposed and evaluated using the benchmark dataset MNIST HAM 10000. To classify images, MobileNet V1 and Inception V3 which are pre-trained CNN models were taken into consideration. The result was given to the web classifier to show the accuracy of the trained model and from the result, they have concluded that they got a better accuracy of 72% produced by Inception V3 whereas MobileNet V1 shows only 58% accuracy.

The classification of melanoma will be critical owing to the factors affecting the classification of the dermatoscopic image samples which basically include presence of noise, air bubbles, hair, similarity between benign and malignant cases and also the low contrast in skin lesion because of the intra-class discrepancy of melanomas. To overcome the above mentioned problems, a Multiple Convolution Neural Network was proposed in [5]. The authors have chosen the ISIC 2016 dataset and to evaluate the model they have used AUC (Area Under ROC). As a result, they got 97.78%, 0.98 AUC for training set and for the testing set they got 85.22% with AUC of 0.81.

In [6], to classify the Atopic Dermatitis Disease which is a multiclass problem will be depending on the value got from Severity Scoring which was calculated using multiclass SVM classifier. The SVM classifier was trained by 22 features which is actually taken from the three main features which are color, texture and redness. There is no specific dataset was taken for the evaluation purpose. The training set was taken from totally 100 images of 55 patients and the testing set was taken from 2 patients. And based on the severity scoring it will be classified into one of the following none, mild, moderate and severe. As the result, they got overall accuracy of 86% with 10-fold cross-validation.

For skin lesion segmentation, a fully convolutional neural network (CNN) was used in [8]. By minimizing a negative multi-label Dice F1 score, implanted feature maps were taken from many intermediate network layers and the negative multi-label Dice F1 score was minimized. By using this, the system got 0.895% AUC value. To separate melanoma from the skin lesion part of the image, a Novel approach using Machine Learning and Information Theory was proposed in [9], in which the authors have compared using division techniques and picture pre-handling methods. From the result, the authors have concluded that Harris Corner Detector and Havrda Entropy provided better sensitivity of 92.45% compared to Otsu and Harris based technology.

The role of AI also has its impact in the field of skin lesions detection. In [10], the skin lesion classification was done using the pre-trained neural network Xception, based on the ABCDE symptoms in which A depicts Asymmetry of the part, B depicts irregularity of the border, C depicts Color of the segmented part, D depicts the Diameter of the lesion and finally E depicts the Evolving nature of the skin lesion. They have used the ISIC dataset and for the image preprocessing they have used the general color constancy algorithm. The system proposed in [11], mainly focusing on skin lesion segmentation. Because of the fuzzy lesion boundary, segmentation will be a difficult problem. The proposed system, alleviate this problem with the help of Iterative Multi-Scale Context-Guided network which is known as MSCGnet.

Segmentation of dermatoscopic images play major role in skin lesion classification. For segmenting skin-lesion classification, an unsupervised algorithm which is called SDI+ is proposed in [14], mainly focusing on to segment dark skin lesions. The process has 3 steps which primarily included the pre-processing, segmentation and post-processing. As a result, for dark skin lesions the better accuracy was achieved.

III. Proposed System

In this paper we are focusing on pre-processing and classification. From the related work, we have observed that many of them were used the imbalanced dataset, so the accuracy is not good. Most of them were not taken the balanced data set, they have used only limited number of images or non-standardized dataset. We have taken the standard HAM10015 dataset which consists of 10015 skin lesion images with 7 categories. The flow diagram of the proposed work is shown in Fig. 1.

A. Preprocessing

To make the dataset balanced, an up-sampling method called Synthetic Minority Oversampling Technique (SMOTE) is used in the proposed method. After up sampling using SMOTE, each type of skin lesion has 6705 samples.

B. Classification

After making the dataset balanced, the next step is to classify the records by training the samples using Machine Learning models. In this paper, we have classified the dataset using most of the machine learning algorithms and compared the algorithms for better Accuracy. The algorithms we have taken here are Naïve Bayes, Logistic Regression, Decision

Tree classifier using Gini index and Entropy, Support Vector Machine using all possible kernel types and XGBoost. We have run the algorithms using default parameters and the algorithm of the proposed work is given below.

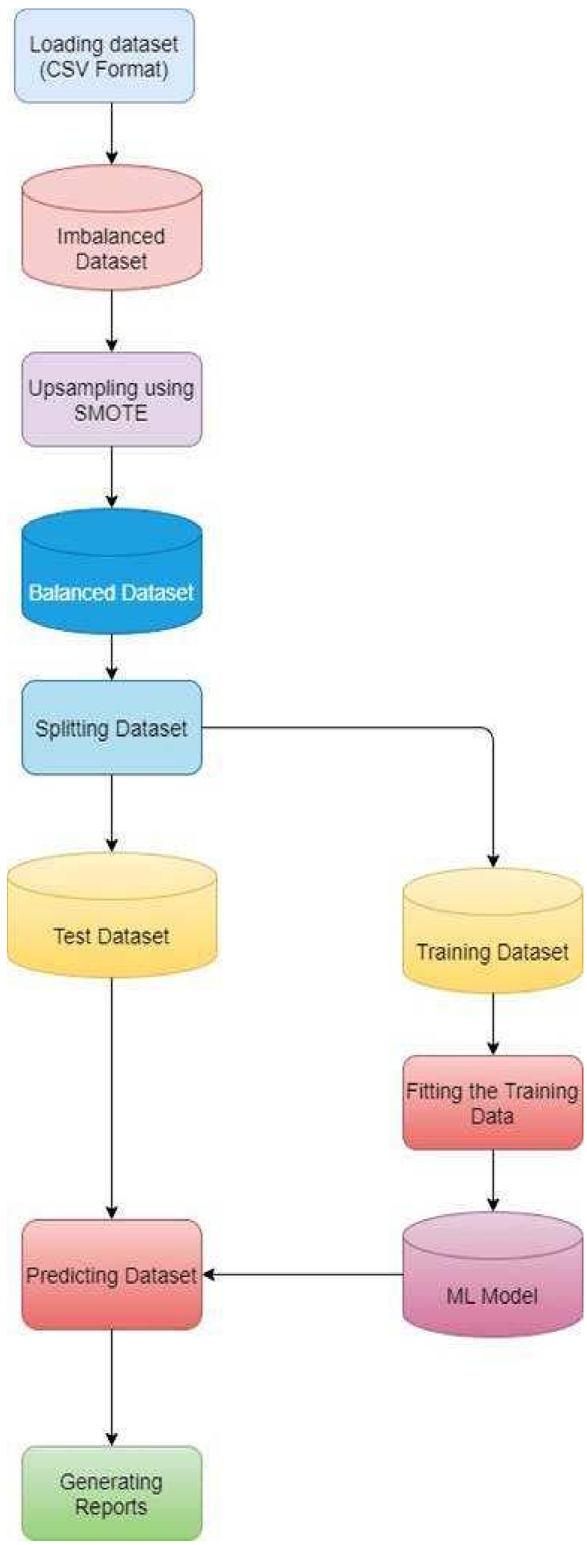

Fig. 1. Flow Diagram of the Proposed System.

C. Algorithm

Loading the dataset

Balancing the dataset using SMOTE

Split the dataset into training and test data

Develop Machine Learning Model like Naïve Bayes, Logistic Regression, Decision Tree classifier using Gini index and Entropy, Support Vector Machine using all possible kernel types and XGBoost using training data set

Evaluate the models using Testing Dataset

IV. Implementation and Results

The proposed system is implemented with Windows 64, 8 GB RAM with Anaconda for Python 3.8 along with Scikit for the implementation of various Machine Learning Algorithms and Pandas library for data manipulation.

MNIST HAM1000 dataset was used for the implementation of the proposed system, which is available through ISIC archive. The dataset consists of 10015 images of 7 classes. Among the entire 10015 records, the dataset has nevus lesions symptom in 6705 samples, dermatofibroma symptom in 114 samples, malignant skin tumors symptom in 1114 samples, benign keratosis symptom in 1099, basal cell carcinoma symptom in 514 samples, actinic keratosis symptom in 327 samples, samples and vascular lesions symptom in 142 records. Since there is a huge difference between the majority and minority classes in the dataset, this is an imbalanced dataset. The Fig. 2 shows the number of samples of each type of skin lesion. In our system, the dataset is in the form of CSV (Comma Separated Value) file, where this file contains the pixel values of images in the dataset. So our CSV file does not contain any features, but actually contains the image in the pixel format (nearly 2351 pixels per images).

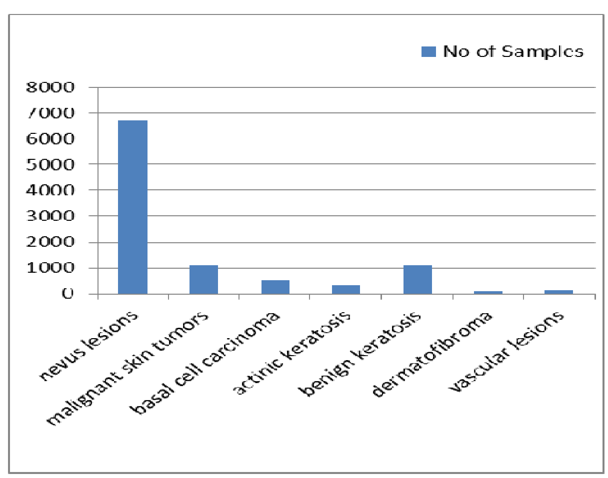

Fig. 2. Number of samples in each type of skin lesions in HAM10000 dataset.

V. Performance Evaluation

To evaluate the model, we have taken Accuracy as the most important evaluation metric. The Accuracy got by different machine learning models in balanced and imbalanced datasets was given in the Table I.

We have also considered f1-score, recall, precision, sensitivity and specificity to evaluate the machine learning models. The results of the above evaluation metrics are represented the Table II.

From Table II, we can found that Support Vector Machine with Polynomial kernel also provides better sensitivity, specificity values and precision, recall, f1-score values compared other algorithms.

From the results, SVM with Polynomial kernel gave a better accuracy of 96.825% with SMOTE compared to other models.

TABLE I
Accuracy Results.

Machine Learning Algorithms	Imbalanced (%)	Balanced (%)
Logistic Regression	67.354	82.451
Naïve Bayes	44.392	39.571
Random Forest	72.013	95.227
Decision tree (Entropy)	63.904	85.447
Decision Tree (Gini)	61.457	83.896
SVM (kernel=Linear)	61.008	95.067
SVM (Kernal=Poly)	**69.695**	**96.825**
SVM (Kernal=rbf)	70.843	92.372
XGBoost	70.43	95.984

TABLE II
Precision, Recall, F1-Score, Sensitivity, Specificity Results.

Machine Learning Algorithm	Precision (%)	Recall (%)	F1-Score (%)	Sensitivity (%)	Specificity (%)
Logistic Regression	82	83	82	83.74	7.19
Naïve Bayes	42	40	40	41.97	88.21
Random Forest	95	95	95	95.82	99.17
Decision Tree (Entropy)	85	86	85	86.03	97.58
Decision Tree (Gini)	85	85	85	85.87	97.59
SVM (Linear)	95	95	95	95.80	99.15
SVM (Polynomial)	**97**	**97**	**97**	**97.29**	**99.45**
SVM (RBF)	92	92	92	92.79	98.73

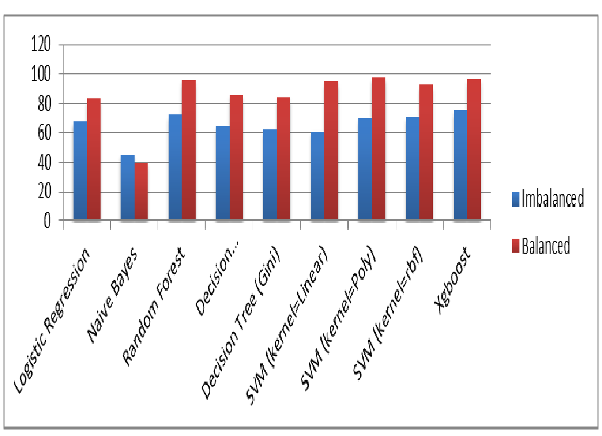

Fig. 3. Comparison of ML models with balanced and imbalanced dataset.

Conclusion

From the results, we have observed that making the dataset balanced will definitely increase the Accuracy. In our paper, the dataset is balanced using an up-sampling with SMOTE. And finally, we have concluded that Support Vector Machine algorithm with Polynomial kernel provides better Accuracy compared to other machine learning algorithms. In our current work, we are just focusing on the pre-processing and classification of dermatoscopic images. In future, we have planned to focus on the segmentation and feature extraction processes of the dermatoscopic images with deep learning models using images itself as input data.

References

[1] X. Wang, X. Jian, H. Ding, and J. Liu, "Bi-directional dermoscopic feature learning and multi-scale consistent decision fusion for skin lesion segmentation," *IEEE Transactions on Image Processing*, vol. 29, pp. 3039–3051, 2020.

[2] C. Kaushik Viknesh, P. Normal Kumar, and R. Seetharaman, "Computer aided diagnostic system for the classification of skin cancer using dermoscopic images," *Proc. 2019 Third International Conference on Inventive Systems and Control (ICISC)*, 2019, pp. 342–345.

[3] S. Krishna and A. S. Remya Ajai, "Analysis of three point checklist and ABCD methods for the feature extraction of dermoscopic images to detect melanoma," *Proc. Ninth International Symposium on Embedded Computing and System Design (ISED)*, 2019, pp. 1–5.

[4] K. E. Purnama, A. K. Hernanda, A. A. Putri Ratnat, I. Nurtanio, A. N. Hidayati, M. H. Purnomo, S. M. Susiki Nugroho, and R. F. Rachmadi, "Disease classification based on dermoscopic skin images using convolutional neural network in teledermatology system," *Proc. International Conference of Computer Engineering, Network, and Intelligent Multimedia (CENIM)*, 2019, pp. 1–5.

[5] Y. Guo, A. S. Ashour, L. Si, and D. P. Mandalaywala, "Multiple convolutional neural network for skin dermoscopic image classification," *Proc. 2018 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, 2018, pp. 365–369.

[6] R. Suhendra, F. Arnia, R. Idroes, N. Earlia, E. Suhartono, "A novel approach to multi-class atopic dermatitis disease severity scoring using multi-class SVM," *Proc. IEEE International Conference on Cybernetics and Computational Intelligence (IEEECYBERNETICSCOM)*, 2019, pp. 35–39.

[7] N. M. Kinyanju, T. Odonga, C. Cintas, N. C. F. Codella, R. Panda, P. Sattigeri, and K. R. Varshney, "Estimating skin tone and effects on classification performance in dermatology datasets," Available online at: https://arxiv.org/abs/1910.13268,2019.

[8] J. Kawahara and G. Hamarneh, "Fully convolutional neural networks to detect clinical dermoscopic features," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 2019, pp. 578–585, 2019.

[9] H. Fabelo, V. Melián, B. Martínez, P. Beltrán, S. Ortega, M. Marrero, G. M. Callicó, R. Sarmiento, I. Castaño, G. Carretero, P. Almeida, A. García, J. A. Hernández, and F. Godtliebsen, "Dermatologic hyperspectral imaging system for skin cancer diagnosis assistance," *Proc. of 2019 XXXIV Conference on Design of Circuits and Integrated Systems (DCIS)*, pp. 1–6, 2019.

[10] D. Gavrilov, L. Lazarenko, and E. Zakirov, "AI recognition in skin pathologies detection" *Proc. 2019 International Conference on Artificial Intelligence: Applications and Innovations (IC-AIAI)*, 2019, pp. 54–542.

[11] Y. Tang, Z. Fang, S. Yuan, C. Zhan, Y. Xing, J. T. Zhou, and F. Yang, "iMSCGnet: Iterative multi-scale context-guided segmentation of skin lesion inDermoscopic images" Published in *IEEE Access*, vol. 8, pp. 39700–39712, 2020.

[12] E. Vocaturo, D. Perna, and E. Zumpano, "Machine learning techniques for automated melanoma detection," *Proc. 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2019, pp. 2310–2317.

[13] A. Kaplan, E. Guldogan, C. Colak, and A. K. Arslan, "Prediction of melanoma from dermoscopic images using deep learning-based artificial intelligence techniques," *2019 International Artificial Intelligence and Data Processing Symposium (IDAP)*, 2019, pp. 1–5.

[14] M. R. Guarracino and L. Maddalena, "SDI+: A novel algorithm for segmenting dermoscopic images," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 2, pp. 481–488, 2019.