Classification of Skin Cancer Dermoscopy Images using Transfer Learning

Haseeb Younis
Department of Computer science
Comsats University Lahore Campus
Lahore, Pakistan
hahhaseeb@gmail.com

Muhammad Hamza Bhatti
Department of Electrical engineering
National University of Computer and Emerging Sciences (FAST NU) Lahore
Lahore, Pakistan
l134449@lhr.nu.edu.pk

Muhammad Azeem
Department of Computer Science
Comsats University Lahore Campus
Lahore,Pakistan
fa17-rcs-022@cuilahore.edu.pk

***Abstract*—Cancer is the second death causing disease in the world. Skin cancer is a major type of cancer and increasing speedily over the past decades. The major causes of skin cancer are Ultraviolet radiation, smoking, DNA mutation and bad lifestyle. Skin cancer is diagnosed by employing clinical screening, dermoscopic analysis, histopathological examination and a biopsy. Early prognosis of skin cancer can help to cure cancer easily and It is a difficult task to diagnosis and classify skin cancer even for expert dermatologists at an early stage. Recent studies have shown that deep learning and transfer learning is very useful in the classification of skin cancer and medical diagnosis. In this paper, we propose an efficient way to classify skin cancer using deep learning. We have tuned the pre-trained MobileNet convolution neural network and trained over the HAM1000 skin lesion dataset. This transfer learning method has shown remarkable categorical accuracy, the weighted average of precision and recall 0.97, 0.90 and 0.91 respectively. This model is lightweight, fast and reliable. It will be helpful for dermatologists to prognosis the skin cancer at its early stage.**

***Index Terms*—Skin cancer classification, skin lesion, Convolution neural networks, Dermoscopy images**

I. Introduction

Cancer is top in death causing diseases and responsible for 9.6 million deaths in 2018 [1], [2]. There are three basic factors that are responsible for cancer: lifestyle, environment and genetic problems. Somatic mutation can be the reason for cancer that is mostly internal cancer and cant be seen with the naked eye or on the skin. Skin cancer is the most common type of cancer that can be caused due to radiation, ultraviolet rays or due to germs. Skin cancer is difficult to detect due to its variant nature. Mostly unusual swelling of tissues and cells can be skin cancer and it is not easy to diagnose that it is cancer or normal swelling. If the diagnosis of cancer is not done on the initial stage it can spread in other organs of the body that is nearly impossible to cure.

There are seven types of skin cancer: Actinic keratoses/ Bowen's disease(akiec), basal cell carcinoma(bcc), benign keratosis-like lesion(bkl), dermatofibroma(df), melanoma(mel), vascular lesion(vasc) and melanocytic nevi(nv). About 86 percent of melanomas and 90 percent of non melanomas skin cancer is caused by ultraviolet radiation [3]. Late diagnosis of these cancer types can damage the other cells of the body and spread in other organs. Skin cancer is very small like the mole at the initial stage, therefore, it is difficult to detect with the naked eye. Dermoscopy has used to diagnose skin cancer at an early stage. It is a noninvasive diagnostic technique for the evaluation of pigmented and non-pigmented lesions on the skin which are not discernible by examination with the naked eye. BCC and BKl are the dullest cancer types that can be cured easily if detected at an early stage. Dermoscopy by inexperienced dermatologists can reduce the accuracy of diagnosis [4]. This implies a need to develop a more accurate and reliable system for the early detection of skin cancer.

The identification of skin cancer is very difficult due to its resemblance. The classification of skin cancer from dermoscopic images can be considered as an image classification problem. The traditional technique of image classification needs prominent features representation that is given to the classifier for training. These features can be texture, color or shape of the color images but in the case of skin cancer, it is difficult to extract the features and to categorize an image base on these features. This draws the attention of researchers to use the deep convolution neural networks(CNNs) to extract the features. Recent studies have proved that CNNs have the ability to extract the high-level (semantics and texture) and low-level features(edges and shapes) automatically due to the abstraction capacity of its layers [5]. Many researchers have work on segmentation, detection and classification of skin cancer using different classification, image processing, machine learning, computer vision and deep learning techniques. Esteva et al [6] and ulzi et al [7] developed the skin cancer classification using CNN. Mohd et al [8] developed the skin cancer classification system for four types of skin cancers using the k-means clustering. Ebtihal et al [9] proposed a classification method for melanoma using the support vector machine. Kawahara et al [10] proposed a multi-tract convolutional neural network to classify ten classes the skin lesion.

These CNNs perform well and extract features automatically but the limitation of CNNs is that these are data-hungry and required huge data for training. To overcome this problem

978-1-7281-5404-6

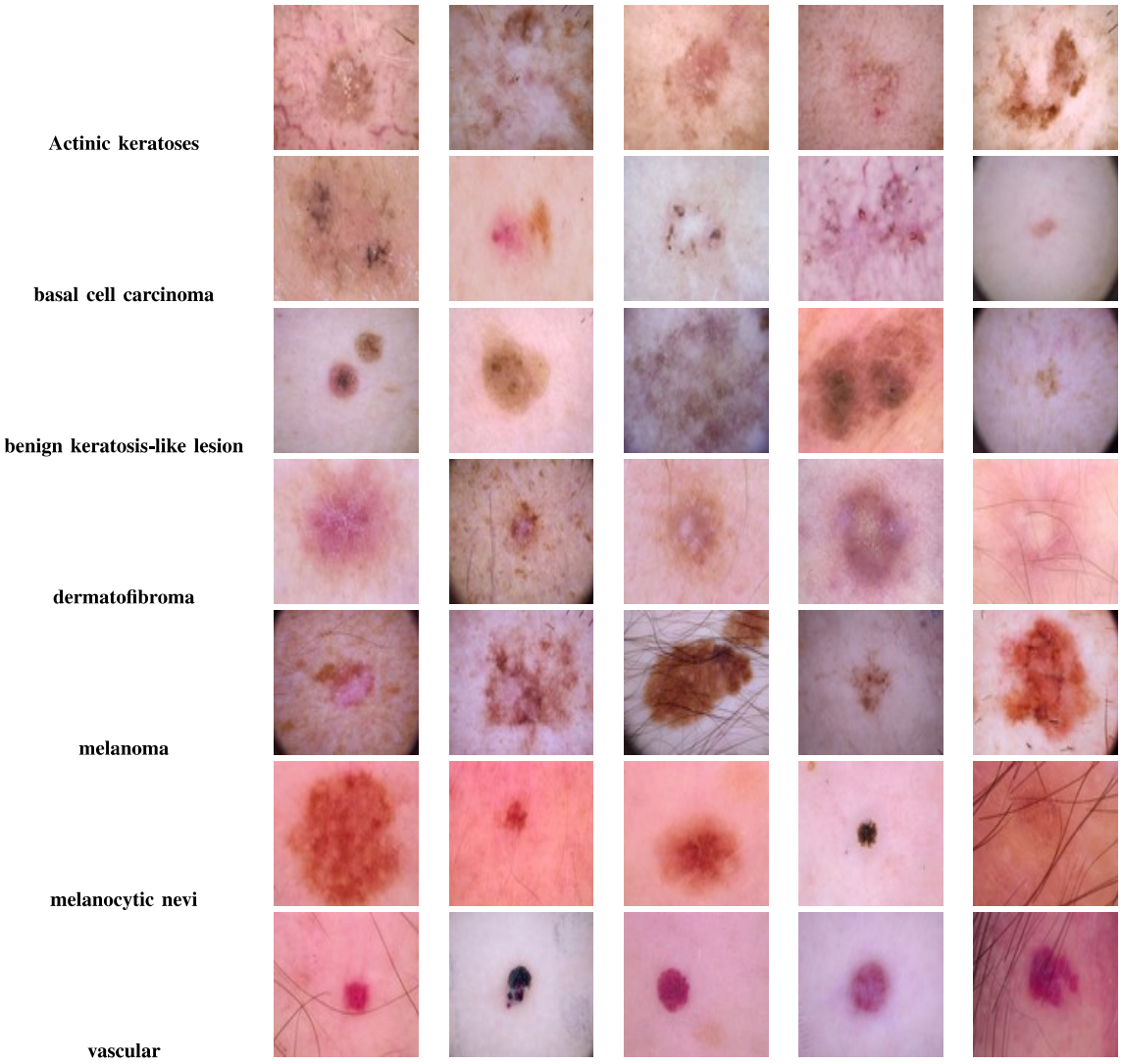

Fig. 1. Image of seven type skin cancers

transfer learning is used [11]. Transfer learning is a simple approach in deep learning that uses initialized weights of pre-trained networks that are trained on other datasets. Anabik et al [12] and Pereira et al [13] used different transfer learning models- ResNet50 [14], DenseNet [15], MobileNet [16] and fine-tuned and ensemble these models to increase the accuracy and robustness of classification but results were not up to mark.

In this paper, we proposed a multi-class skin cancer classification method to classify dermoscopic skin cancer images into one of the seven cancer classes- akiec, bcc, mel, bkl, df, vasc and nv. We find-tune the MobileNet convolutional neural network [16] pre-trained on 1.3 million images and train on the Harvard HAM1000 skin cancer dataset [17]. that contains 10015 dermatoscopic images. We also analyzed the dataset to discover the relation of skin lesions with different parameters to increase the understanding of skin cancer.

II. Methods

A. Dataset

In this paper, we used the HAM10000 dataset[17], a large collection of multi-source dermatoscopic images of common pigmented skin lesions acquired and stored by different modalities and from different population. The dataset contains 10015 images of seven skin cancer type that includes 6705 melanocytic nevi (nv), 1113 melanoma (mel), 1099 benign keratosis-like lesion (bkl), 514 basal cell carcinoma (bcc), 327 Actinic keratoses/ Bowen's disease (akiec), 142 vascular(vasc) and 115 dermatofibroma (df) images. Sample images of all seven skin cancer types are shown in Figure 1.

B. Data preprocessing and training algorithm

We split our dataset to training data (9000 images) and validation data to authentic our model. The original resolution of images was 600x450 pixels that downscaled to 224x224 pixels to make an image compatible with our training model

TABLE I: Weighted average precision report for seven classes of skin cancer

Classes	Precision
akeic	50
bcc	55
bkl	77
df	20
mel	28
nv	97
vasc	77
Weight Average	0.90

MobileNet. Due to the unbalance distribution of data it was difficult to train the model accurately and the loss rate was increasing. Data augmentation[18] is a very useful technique in such cases to balance the unbalanced data. We used augmented augmentation for the minority classes to balance the whole data. After balancing our data we import and fine-tune the MobileNet training model to train on our dataset. MobileNet have 93 layers in total, we dropped the last 5 layers of and include and dense layer with softmax activation function for prediction. Then we freeze the weights of all layers except the last 25 layers that will be trained. We used a 70/30 train/test spilt of the dataset and 30 epochs. The Adam optimizer, Categorical crossentropy and accuracy were used to evaluate the performance of the MobileNet network.

C. Evaluation Measures

We used Accuracy (Acc), precision and Recall (TPR) for the evaluation of our model

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

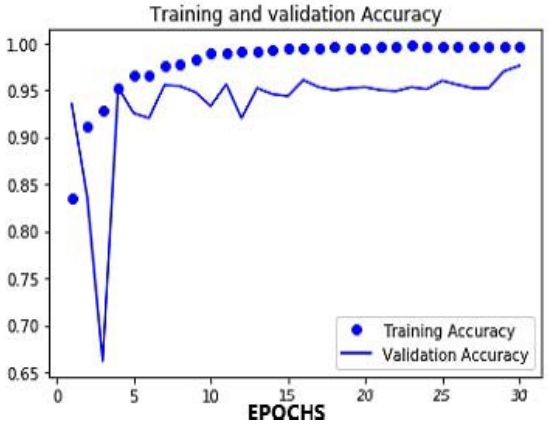

Fig. 2: Categorical Accuracy of Model

TABLE II: Implementation results comparison between previous research works

	Accuracy%	Recall%
Esteva[6]	72.1	96
Mohd [8]	83.33	-
Ebtihal[9]	90.32	93.97
Pereira[13]	91.5	-
Ulji[7]	94.2	90.74
Anabik[12]	77.5	-
Proposed system	**97.07**	**91.34**

III. Results and Discussion

The model was train on core i7-7800 with 16 GB ram and 11 GB GPU (Gforce 1080-TI). Model evaluation was performed by calculating the training and validation accuracy, loss and accuracy curve for the validation of model performance, and comparison with recent studies.

A. Model Validation

The model validation was conducted on 1000 sample images from the validation dataset. To understand the generalized performance of the model we calculated the weighted average precision. This weighted average precision was calculated for the seven classes of skin cancer that was 0.90. The detailed seven class classification report is shown in Table 1. We also have compared our model performance with recent studies like Esteve et al [6], Mohd et al [8], Ebtihal et al [9], Pereira et al [13], Ulji et al [7] and Anabik et al [12] shown in table 2. In these studies most of them have worked using three or four classes, Esteve [6] reported 72.1 percent classification accuracy, Mohd [8] reported 83.33 percent classification accuracy, Ebithal [8] reported 90.32 percent accuracy, Pereira et al [13], Ulji et al [7] reported 91.5 and 94.2 percent accuracy respectively. From all these competitors anabik et al [12] have used the transfer learning and his accuracy is 77.5 on seven classes. We achieved the highest categorical accuracy 97.07 further our model is lightweight that can be employ easily and predict the class approximately within 2-3 seconds.

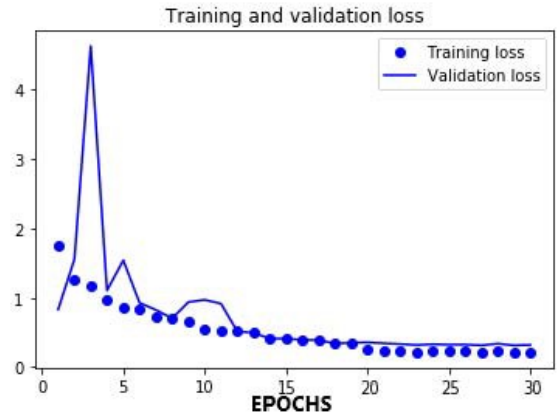

Fig. 3: Validation and Training Loss

B. Accuracy and loss curves

We have draw loss and accuracy curve to evaluate the performance of our model. These curves show that accuracy is increasing and loss is decreasing as the numbers of iterations are increasing. The gap between validation and training curves shows good fitting and the model will perform well on unseen images. The accuracy curve is shown in Figure 2 and the loss curve is shown in Figure 3.

IV. Conclusion

In this paper, we have classified seven types of skin cancers using a pre-trained MobileNet convolutional neural network with high accuracy. Skin cancer is increasing over the several past decades and diagnosis is difficult and challenging even for expert dermatologists due to variability in appearance. In recent studies, deep learning is playing a remarkable role in the classification of skin cancer, visual recognition, and medical diagnosis. Here we present skin cancer classification method using deep learning over HAM1000 skin lesion dataset having seven skin cancer classes: Actinic keratoses/ Bowen's disease (akiec), basal cell carcinoma (bcc), benign keratosis like lesion (bkl), dermatofibroma (df), melanoma (mel), vascular lesion (vasc) and melanocytic nevi (nv). The model shows the categorical accuracy of 97 percent and weighted average precision and recall 0.90 and 0.91 respectively. MobileNet is a lightweight, fast, robust and accurate model. In our case, It classifies the skin cancer within 2-3 seconds with remarkable accuracy.

References

[1] M. Plummer, C. de Martel, J. Vignat, J. Ferlay, F. Bray, and S. Franceschi, "Global burden of cancers attributable to infections in 2012: a synthetic analysis," Lancet Glob. Heal., vol. 4, no. 9, pp. e609–e616, Sep. 2016.

[2] "Cancer Statistics - National Cancer Institute." [Online]. Available: https://www.cancer.gov/about-cancer/understanding/statistics. [Accessed: 30-Aug-2019].

[3] D. M. Parkin, D. Mesher, and P. Sasieni, "13. Cancers attributable to solar (ultraviolet) radiation exposure in the UK in 2010," Br. J. Cancer, vol. 105, no. S2, pp. S66–S69, Dec. 2011.

[4] H. Kittler, H. Pehamberger, K. Wolff, and M. Binder, "Diagnostic accuracy of dermoscopy," Lancet Oncol., vol. 3, no. 3, pp. 159–165, Mar. 2002.

[5] F. Xie, H. Fan, Y. Li, Z. Jiang, R. Meng, and A. Bovik, "Melanoma Classification on Dermoscopy Images Using a Neural Network Ensemble Model," IEEE Trans. Med. Imaging, vol. 36, no. 3, pp. 849–858, Mar. 2017.

[6] A. Esteva et al., "Dermatologist-level classification of skin cancer with deep neural networks," Nature, vol. 542, no. 7639, pp. 115–118, Feb. 2017.

[7] U.-O. Dorj, K.-K. Lee, J.-Y. Choi, and M. Lee, "The skin cancer classification using deep convolutional neural network," Multimed. Tools Appl., vol. 77, no. 8, pp. 9909–9924, Apr. 2018.

[8] Anas, Mohd, Kailash Gupta, and Shafeeq Ahmad. "Skin cancer classification using K-means clustering." International Journal of Technical Research and Applications 5.1 (2017): 62-65.

[9] Almansour, Ebtihal, and M. Arfan Jaffar. "Classification of Dermoscopic skin cancer images using color and hybrid texture features." IJCSNS Int J Comput Sci Netw Secur 16.4 (2016): 135-9.

[10] J. Kawahara and G. Hamarneh, "Multi-resolution-Tract CNN with Hybrid Pretrained and Skin-Lesion Trained Layers," Springer, Cham, 2016, pp. 164–171.

[11] A. Pal, A. Chaturvedi, U. Garain, A. Chandra, and R. Chatterjee, "Severity grading of psoriatic plaques using deep CNN based multi-task learning," in 2016 23rd International Conference on Pattern Recognition (ICPR), 2016, pp. 1478–1483.

[12] Pal, Anabik, Sounak Ray, and Utpal Garain. "Skin disease identification from dermoscopy images using deep convolutional neural network." arXiv preprint arXiv:1807.09163 (2018).

[13] F. Pereira dos Santos and M. Antonelli Ponti, "Robust Feature Spaces from Pre-Trained Deep Network Layers for Skin Lesion Classification," in 2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), 2018, pp. 189–196.

[14] He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778. 2016.

[15] Huang, Gao, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q. Weinberger. "Densely connected convolutional networks." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4700-4708. 2017.

[16] Howard, Andrew G., Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. "Mobilenets: Efficient convolutional neural networks for mobile vision applications." arXiv preprint arXiv:1704.04861 (2017).

[17] P. Tschandl, C. Rosendahl, and H. Kittler, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," Sci. Data, vol. 5, no. 1, p. 180161, Dec. 2018.

[18] A. Mikolajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," in 2018 International Interdisciplinary PhD Workshop (IIPhDW), 2018, pp. 117–122.