Performance Evaluation of Novel Convolution Neural Network Architecture for Melanoma Skin Cancer Diagnosis on Different Hardware Processing Units

Mohammad Naved Qureshi\(^1\) and Mohammad Sarosh Umar\(^2\)

\(^1\)Assistant Professor, EES, University Polytechnic, Faculty of Engineering & Technology, AMU, Aligarh, India
\(^2\)Professor, Department of Computer Engineering, ZHCET, Faculty of Engineering & Technology, AMU, Aligarh, India.

E-mail: navedmohd786@gmail.com, saroshumar@zhcet.ac.in

Abstract. Skin cancer is one of the major public health concerns among the white population with more than a hundred thousand cases every year. Melanoma is one of the deadliest forms of skin cancer which is responsible for thousands of deaths in US alone in recent years and, therefore, early diagnosis is very important to increase the survival rate of melanoma patients. In last few years Deep neural networks have been utilized by researchers to build best models for classifying or diagnosing skin cancer. In this paper Deep neural network-based CNN architectures to classify Melanoma is proposed. The CNN architecture proposed in this work is implemented on CPU, GPU and TPU and the performance of the model is shown on all these platforms. The proposed model is compared to other works done so far for melanoma diagnosis in terms of various performance metrics like prediction accuracy, specificity, sensitivity and it is observed that the proposed models outperformed them. The dataset utilized in training and testing the proposed models is ISIC archive dataset which contains 4750 skin images for two classes i.e. melanoma and benign. The results of our study have proved that utilizing GPU and TPU speeds up the training 38 times faster than CPU and can accelerate the performance of CNN for features extraction, optimization and classification of skin cancer images and the proposed model has outperformed the other models compared with it.

1. Introduction

Among all types of cancer skin cancer is most common. Taking about melanoma than it accounts for only 1% of all skin cancers but it is responsible for large number of casualties or deaths. According to the survey of American Cancer Society total number of melanoma cases estimated to be diagnosed in the United States alone for 2020 are about 100,350 and number of people who have died of melanoma in 2020 is estimated to be about 6,850[1]. If we consider skin tone than white peoples‘ have chances of melanoma 20 times more than in African Americans. Overall, the risk of getting melanoma for white people once in their lifetime is about 2.6% (1 in 38), for blacks it is 0.1% (1 in 1,000), and for Hispanics it is 0.6% (1 in 167) [1]. Nowadays, lot of advancement is done in in the field of medical treatment techniques, such as radiation therapy, immunotherapy and surgery combined with both [2]. The survival rate is 5 years‘ if melanoma reaches advanced stage which is around 15% while more than 95% if it is early-stage melanoma [3]. This difference in survival rate in early and later stage shows that there is a need for early diagnosis of melanoma for the survival of the patients. Diagnosis of melanoma by physicians often rely on years of experience and evaluation of patient’s lesions on a case-by-case basis. The patient’s local lesion patterns are examined by the dermatologist compared to that of the entire body...
Without computer-based assistance, the accuracy of clinical diagnosis for detection of melanoma is found to be between 65 to 80% [5]. Nowadays, computer-aided decision systems are used when evaluating and diagnosing medical images accurately [6]. For example, computer aided diagnosis (CAD) system for detecting breast cancer on mammograms in the US [7]. It is one of the major research subjects in medical imaging and diagnostic radiology [7]. For early detection of a disease an accurate CAD system can be used thereby allowing for earlier treatment, which could save lives [8]. In 1990 CAD systems for skin lesion detection were introduced [9]. To address the challenging task of skin classification a lot of methods have been published. Most of the algorithms proposed used the manual evaluation method which is based on the ABCD rules proposed by Nachbar et al. in [10]. The rules are based on asymmetry (A), border (B), color (C), and differential structure (D). For the proper integration of these rules into a computer-based system several problems should be handled. The first is the segmentation of the skin lesion which is very important for the analysis of asymmetry and the border. Many methods are proposed for skin segmentation like thresholding technique, region and edge based techniques, soft computing techniques and deformable models [11] and one proposed by [12] using neutrosophic logic. Most of the techniques including traditional machine learning techniques developed for skin cancer detection are dependent on hand crafted features. Tasks like pre-processing of image, image segmentation and feature extraction are very complex in basic machine learning approaches which leads to degradation of system performance regarding training time and classification accuracy. To overcome all of these shortcomings of basic machine learning techniques, deep learning concept came into existence which is been used to extract the important information or features directly from the input raw images during training phase and then use those features for classifications process [13, 14]. So there is no need for feature extraction separately. Deep learning based on Convolution Neural Network (CNN) models has achieved great success in the last few years in the field of medical image analysis like detection of diabetic retinopathy, breast cancer [15], tumor detection [16], skin disease and its classification [17] etc. Provided the large datasets the CNN application works very well but if the dataset is small then it fails to achieve significant performance gains. To achieve higher accuracy on small data set transfer learning technique can be used [18]. Instead of random initialization of weights a pre-trained network on large datasets, i.e. ImageNet [19], that contains 1.2 million images labeled with 1,000 classes can be used.

In this paper, we have developed a novel CNN model for melanoma classification and have implemented it on CPU, GPU and TPU. The model is trained using variable learning rate and tested on skin lesion ISIC Archive dataset and the results are compared with some other work done so far in melanoma diagnosis. The rest of the paper is structured as follows: Section 2 reviews some related work, Section 3 describes the materials and methods, Section 4 presents the results of experiments and finally, Section 5 offers concluding remarks and directions for future work.

2. Related Work
Recently, deep learning techniques for image classification have become very popular since the publication of AlexNet architecture in 2012. In this section, focus is on the recent studies that have used deep learning for the skin cancer classification. A model proposed in [20] focused on knowledge transfer by utilizing pretrained network which was trained don image net datasets. The authors have shown the significance of more deep models for classification task. In [21] an automated deep learning based melanoma recognition system was proposed, combined with so called hand-crafted RSurf features and Local Binary Patterns. Their results show high classification accuracy compared to some other works. In the system proposed in [22], preprocessing of input images are done in order to reduce noise and illumination and fed to a pre-trained deep learning based convolutional neural network. The CNN classifier proposed used binary classification to classify melanoma and benign cases. A deep learning based approach is used in [23] for automatic segmentation of skin tumor from the input skin image using fully convolutional deconvolutional architecture and a simple CNN and VGG-16 architecture used for the classification of skin tumor. The work done in [24] evaluate the main architectures of CNN for melanoma skin cancer diagnosis. The PH2 dataset of dermatoscopic images is used on different
classifiers and the best result was obtained by the combination of the VGG19 network and logistic regression having accuracy of 92.5% and 85.71% of precision. Another deep learning based ensemble model is proposed in [25] using aggregation of robust CNNs into one single neural net architecture, where the weighted output of the member CNNs provide the final classification.

A similar method is proposed in [26], just like [25], where the ensemble of deep CNN is used for ternary classification i.e. melanoma, nevus, and seborrheic keratosis. A deep neural network architecture is developed by fusing the outputs of the classification layers of four different models. The weighted output of the member CNNs fused together gives the final classification. Several methods are proposed for fusing the different architectures and each of the fusion method outperforms the individual network architecture in terms of classification accuracy. The average value of area under the receiver operating characteristic curve is 0.891 for ternary classification task. Similarly, a method proposed in [27] consists of many CNNs each of which is having different architecture and each of them consist of a number of pre-trained networks having same architecture with different hyper parameter settings and are fine-tuned on skin lesion images. Different SVM classifiers were trained from features obtained from the deep layers of each network. Finally, the classification is obtained by taking the average value of prediction probability which was obtained by fusion of classification vectors from different sets. The area under receiver operating characteristic curve of 87.3% for melanoma classification is obtained for propose algorithm. In the method proposed in [28] a fully convolutional deconvolutional architecture is used for skin tumor segmentation from the input skin image. For classification, concept of transfer learning is used with a simple CNN and VGG-16 architecture. The average value of 0.507 is obtained for Jaccard index and the AUC of 0.763 and 0.869 for two different lesion classifications using VGG16 is obtained using transfer learning approach. The work done in [29] had tested different classifiers on the PH2 dataset of dermatoscopic images by extraction of the characteristic vectors by the networks. The best performance is achieved when the VGG19 network is combined with logistic regression, obtaining the accuracy of 92.5% and 85.71% of precision.

3. Materials and methods
This study aims to perform a binary classification of melanoma versus benign from skin images. The sequence of important general steps in the proposed skin lesion classification method are as follows:

a. Data Preparation
b. Selecting the CNN model architecture
c. Training the model to build a classifier and predicting labels for test data

3.1. Dataset Preparation
Before training the classifier the dataset is passed through various stages as follows:

3.1.1. Preprocessing and Data Splitting
In our proposed method, the pre-processing steps are kept to a minimum so that our models have better generalization ability when tested on other datasets. We have used International Skin Imaging Collaboration (ISIC) archive dataset, which consists of 4570 images of two class benign and malignant. To prepare the images for the CNN, each of the training images was resized to 224 pixels by 224 pixels. Then, the input images are rescaled by normalizing the pixel values to a [0, 1] range by multiplying each pixel value by a factor of 1/255. After that data is split into train, validation, and test in the ratio 60:30:10 respectively.

We get the images in each set of data in the form of matrices of following size:
Train data size: (2755, 224, 224, 3)
Validation data size: (1358, 224, 224, 3)
Test data size: (457, 224, 224, 3)
3.1.2. Data Augmentation
To avoid the overfitting of the proposed CNN models, data augmentation of the training dataset is done. The data can be augmented in lot of ways, and for this study we have utilized vertical flipping, horizontal flipping, rotating images with 180° angle, width shift, height shift and zooming by 0.15, 0.15 and 0.25 respectively. Horizontal and vertical flip are also done for each image. Using these transformations, we have not increased our data size and it remains same as the original one.

3.2. Convolution Neural Network Architecture
The CNN proposed for classification of melanoma and benign is shown in the figure 1 below. This CNN has very powerful convolve filters for getting important features from the input images. The model consists of 3 convolution layer each of them followed by max pool layer. The first convolution layer has 16 filters of 3x3, second has 32 filters of 3x3 and last convolution layer has 64 filters of 3x3 each. Each of the convolution layer used relu as an activation function. Each of the max pooling layer has stride of 2x2 with no zero padding so as to downgrade the size of the input image. A drop out of 50% is added after max pool, which is actually a form of regularization and its purpose is to prevent overfitting by providing enhanced testing accuracy. For each batch in our training set, dropout layers, with probability \( p \), randomly disconnect inputs from the preceding layer to the next layer in the network architecture. Following this the Flatten layer is there which takes multi-dimensional volume and flattens it into a 1D array prior to feeding the inputs into the next layer i.e. Dense or fully-connected layer. The dense layer has 512 output nodes after which another drop out of 50% is applied and finally the last fully connected or output layer is there with two nodes for each of the two classes in the dataset.

![Figure 1. Proposed CNN Model](image-url)
3.3. Training the model to build a classifier and predicting labels for test data
In this section the training and testing of the proposed model is discussed. The model is trained on ISIC dataset of skin lesion images which is publicly available [28]. The training dataset consists of 2755 images of benign and malignant and the validation is also done side by side on the set of 1358 images. The model is tested on 457 images obtained after partitioning of data above. The model is fine tuned with various hyper parameters as discussed below:

3.3.1 Model Hyper Parameters:
Hyper parameters are properties that control the entire training process. The proposed model hyper parameters are:

3.3.1.1 Learning Rate:
When training deep convolution neural networks, the learning rate has to be fine-tuned as the training progresses. We have used the variable learning rate which increase linearly in the beginning and then exponentially decays as the training progresses as shown in the figure 2. The method we used for variable learning rate (lr) is as follows:

\[
\text{if current\_epoch < power\_epoch then:} \\
\begin{align*}
    lr &= \left( \frac{\text{max\_lr} - \text{starting\_lr}}{\text{power\_epoch}} \right)^{\text{current\_epoch}} + \text{starting\_lr} \\
    \text{else:} \\
    lr &= \text{max\_lr} \times \text{decay\_rate}^{\frac{\text{current\_epoch} - \text{power\_epoch}}{\text{power\_epoch}}} \\
\end{align*}
\]

here, \( \text{decay\_rate} = \left( \frac{\text{min\_lr}}{\text{max\_lr}} \right)^{\frac{\text{global\_epoch}}{\text{power\_epoch}}} \)

Where,

- \( \text{global\_learning\_rate} = 1 \times 10^{-4} \) or 0.0001
- \( \text{global\_epoch} = 64 \); (i.e. number of times learning model algorithm has to train on entiredataset)
- \( \text{min\_lr} = 1 \times 10^{-6} \) or 0.000001; (i.e. minimum learning rate).
- \( \text{max\_lr} = \text{global\_learning\_rate} \) (i.e. maximum learning rate)

\( \text{starting\_lr} = \left( \frac{\text{global\_learning\_rate}}{\text{total number of processing unit threads}} \right) \); (i.e. starting learning rate)

here, number of processing threads for CPU or GPU is 2 and for TPU is 8.

- \( \text{current\_epoch} = \) the epoch running currently during training process (i.e. from 0 to 63).
- \( \text{power\_epoch} = \text{global\_epoch}/20 \); (i.e. taking only integer part of the division).
3.3.1.2 Optimizer
We have used Adam [29] optimizer which is designed specifically for training deep convolution neural networks and it is an adaptive learning rate optimization algorithm.

3.3.1.3 Loss function
A loss function quantifies the difference between our predicted class labels with ground-truth labels. We have used sparse categorical cross entropy loss to compute the cross entropy loss between the labels and predictions.

3.3.1.4 Activation function
Activation function are used in hidden layer as well as output layer of CNN. The commonly used activation functions in hidden layers are ReLu, sigmoid and tanh etc. But we’re here to talk about the activation function in the output layer. A function that takes the values from output nodes and transforms them into a probability distribution is softmax activation function. It reports back the confidence score for each class in the dataset. The total value of class score will sum up to one as we are dealing with probabilities in softmax output. The predicted class (benign or malignant in our dataset) is the one with the highest confidence score.

3.3.1.5 Learning Process Controller
We have use the early stopping criterion for controlling the learning process of the classifier. The classifier will stop learning when the validation loss will not decrease further beyond a certain point. So in this case the number of epochs may or may not reach to total epochs decided for training the classifier model. This will help us to reduce unnecessary training of our model as we may get best performance in less epochs also.

3.3.1.6 Epoch and Batch Size
One training epoch means that one pass has been made through the training dataset by the learning algorithm. In our training of model, we have used total of 64 epochs but number of epochs may not reach 64 because of early stopping criterion used. Batch size tells about the number of training examples used to estimate the error gradient. We have used the batch size of 32 during the training of our classifier.

Figure 2. Variable Learning rate graph for CPU, GPU and TPU
model. This means that before the model weights are updated 32 samples from the training dataset will be used to estimate the error gradient.

### 3.3.2 Testing of model

After training is completed the CNN classifier model is tested on unlabeled dataset to predict whether the image is benign or malignant. The test set consists of total of 457 images belonging to two classes. The loss function used is sparse categorical cross entropy loss and various metrics are calculated to evaluate the performance of the classifier.

### 4. Experimental Results

In this section the proposed method is evaluated using the commonly used metrics on training, validation and test dataset. The proposed model is implemented on different processing units like CPU, GPU and TPU in google colabs. The result of the model performance is evaluated on all these processing units using the same dataset and same model hyper parameters. We have divided this section into two parts i.e. training/validation results and testing results as follows:

#### 4.1 Training & validation results

The training and validations results of the proposed model on CPU, GPU and TPU is shown in the table 1 below and the graph is also shown in the figure 3, 4, 5 and 6.

From the table 1 and figure 3,4,5 and 6 it is observed that model performance is better on GPU as compared to TPU and CPU. All the evaluation metrics have higher values in case of GPU implementation. Training accuracy and time are very less on GPU implementation. It is also observed that due to the use of variable learning rate the number of epochs has been reduced in all the three implementations. The model only learns up to when it reaches maximum value of validation accuracy and then it stops when validation accuracy is not increased beyond a certain epoch. Thus a lot of time is saved during model training as we don’t have to go up to 64 epochs.

| Evaluation Metrics | Processing Unit                  |
|---------------------|----------------------------------|
|                     | CPU (Intel(R) Xeon(R) CPU @ 2.20GHz) | GPU (Tesla P100-PCIE-16GB) | TPU (Cores: 8) |
| Time Taken          | 2817 seconds                      | 75 seconds                  | 95.42 seconds  |
| Epochs Reached      | 18                               | 20                         | 15             |
| Training Loss       | 0.1148                            | 0.1098                      | 0.1304         |
| Training Accuracy   | 95.83                             | 95.93                       | 95.39          |
| Validation Loss     | 0.1696                            | 0.1739                      | 0.1772         |
| Validation Accuracy | 94.11                             | 94.32                       | 94.26          |
Figure 3. Training accuracy CPU vs GPU vs TPU

Figure 4. Validation Accuracy CPU vs GPU vs TPU
Figure 5. Training loss CPU vs GPU vs TPU

Figure 6. Validation loss CPU vs GPU vs TPU
4.2 Test results

The proposed model is tested on set of 457 images belonging to two categories benign and malignant. The loss function used here is sparse categorical cross entropy loss and activation at output layer to predict class score is softmax activation function. The various metrics results are shown in the table 2 below.

| Evaluation Metrics | CPU  | GPU   | TPU  |
|--------------------|------|-------|------|
| sparse_cc_loss     | 0.170001 | 0.166274 | 0.16216 |
| benign_accuracy    | 94.52 | 95.19 | 94.97 |
| malignant_accuracy | 94.53 | 95.18 | 94.96 |
| avg_accuracy       | 94.53 | 95.19 | 94.97 |
| avg_AUC            | 97.67 | 97.88 | 98.01 |
| avg_precision      | 0.914167 | 0.925434 | 0.911741 |
| avg_recall         | 0.914745 | 0.926334 | 0.912533 |
| avg_specificity    | 0.914745 | 0.926334 | 0.912533 |
| avg_sensitivity    | 0.914745 | 0.926334 | 0.912533 |

From the above results we can say that the proposed model has slightly higher accuracy on GPU and higher average AUC on TPU. Others metrics like precision, recall, sensitivity and specificity is more on GPU than TPU and CPU. Now in the figures below ground truth vs predicted images are shown.

Figure 7. Ground truth vs Predicted image on CPU
Now the proposed model is compared with some other work done so far on same dataset that we have used. The table 3 below shows the result and it’s clear that our proposed model has outperformed the other works done so far in this field in terms of sensitivity, specificity and accuracy.

| Methods                                      | Sensitivity | Specificity | Accuracy |
|----------------------------------------------|-------------|-------------|----------|
| Spotmole [31]                                | 0.82        | 0.57        | 0.67     |
| MED-NODE texture descriptor [32]             | 0.62        | 0.85        | 0.76     |
| [30]                                         | 0.81        | 0.80        | 0.81     |
| [33]                                         | 0.46        | 0.87        | 0.70     |
| Proposed CNN model                           | **0.92**    | **0.92**    | **0.98** |

5. Conclusion & Future Work
In this paper, a novel CNN model based on deep learning is proposed for melanoma classification with minimum requirements for preprocessing operations on the skin images. The proposed method is implemented on CPU, GPU and TPU in google colabs. Early stopping criterion and variable learning rate used in this work has improved the model performance specially on GPU as the time required for
training on GPU and TPU is almost 38 times faster as compared to CPU. The general performance of the proposed method is far better than other state-of-the-art methods. The concept of transfer learning can be utilized in future in which a pretrained model on imagenet dataset can be used for skin classification task by fine tuning the model and adding or removing some layers in the network depending upon the problem domain and number of classes in the dataset. The same can be implemented on different processing units to get better results. Image segmentation can also be done as a part of preprocessing to get better output results.

References

[1] American cancer society Available at https://www.cancer.org/cancer/melanoma-skin-cancer.html

[2] Barker CA, Postow MA (2014) Combinations of radiation therapy and immuno therapy for melanoma: A review of clinical outcomes J. radia t. Oncol. Biol. Phys. vol. 88, no. 5pp. 986–997.

[3] Balch CM (2009) Final version of 2009 AJCC melanoma staging and classification J. Clin. Oncol vol. 27 no. 36 pp. 6199–6206.

[4] Gachon J, Beaulieu P, Sei JF, Gouernet J, Claudel JP, Lemaître M, Richard MA, and Grob JJ (2005) First prospective study of the recognition process of melanoma in dermatological practice Arch. of Dermatol 141(4) pp 434–8.

[5] Argenziano G, Soyer HP (2001) Dermoscopy of pigmented skin lesions a valuable tool for early diagnosis of melanoma. The Lancet Oncol 2(7) pp. 443–9.

[6] Masood A, Ali-Jumaily A (2013) Computer aided diagnostic support system for skin cancer: a review of techniques and algorithms Int. J. of biomed. imaging.

[7] Doi K (2007) Computer-aided diagnosis in medical imaging: historical review, current status and future potential. Computerized medical imaging and graphics. 31(4-5), pp. 198–211.

[8] Abdel-Zaher AM, Eldeib AM. (2016) Breast cancer classification using deep belief networks. Expert syst. with appl. 46. pp. 139–144.

[9] White R, Rigel DS, Friedman R (1992) Computer applications in the diagnosis and prognosis of malignant melanoma, Dermatol. Clin. 9(4) pp. 695–702.

[10] Nachbar F, Stolz W, Mertke1T, Cognetta AB, Vogt T, Landthaler M, Plewig G, (1994) The ABCD rule of dermatoscopic high prospective value in the diagnosis of doubtful melanocytic skin lesions. J. Am Acad. Dermatol. 30 (4) pp. 551–559.

[11] Kaur, Dilpreet and Yadwinder Kaur. (2014), Various Image Segmentation Techniques: A Review, Int. J. of Comp. Sci. and mob. Comput. Vol.3 Issue.5 pg 809-814.

[12] Qureshi MN, Ahamad MV (2018) An improved method for image segmentation using K-means clustering with neutrosophic logic. Procedia comp. sci. 132, pp. 534-40.

[13] LeCun Y, Bengio Y, Hinton V (2015) Deep learning. Nature 521 (7553) pp. 436.

[14] Bengio Y, Courville A, Vincent, P (2013), Representation learning: A review and new perspectives IEEE trans. on pattern anal. and mach. Int e l. 35 (8) pp.1798–1828.

[15] Shen, Li &Margolies, Laurie & Rothstein, Joseph &Fluder, Eugene & McBride, Russell &Sth, Weiva. (2019). Deep Learning to Improve Breast Cancer Detection on Screening Mammography. Scientific Reports. 9. 1-12. 10.1038/s41598-019-48995-4.

[16] Cruz-Roa AA, Ovalle JEA, Madabushi A, Osorio FAG (2013) A deep learning architecture for image representation, visual interpretability and automated basal-cell carcinoma cancer detection, Int. conf. on med. Image comput. and comp. assist. Interv.Springer, pp. 403–410.

[17] Esteva A, Kuprel B, Tblum S, Deep networks for early stage skin disease and skin cancer classification Project Report Stanford University.

[18] Pan, SinnoJalin and Qiang Yang 2010, A Survey on Transfer Learning IEEE trans. on knowl. and data engg. 22 pp1345-1359.

[19] Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla V, Bernstein M, and Berg AC2015 Imagenet large scale visual recognition challenge, Int. j. of comp. vision, 115(3).

[20] Menegola, Afonso et al. 2017 Knowledge transfer for melanoma screening with deep learning. IEEE 14th Int. symp. on biomed. Imaging (ISBI 2017) (2017): n. pag. Crossref. Web.

[21] Majtner V, Yildirim-Yayilgan S, Hardeberg JY 2016 Combining deep learning and hand-crafted features for skin lesion classification Sixth Int. conf. on Image process. Theory, Tools and appl. (IPTA)Oulu, pp. 1-6, doi: 10.1109/IPTA.2016.7821017.

[22] E. Nasr-Esfahani et al.2016 Melanoma detection by analysis of clinical images using convolutional neural network.38th Annual Int. conf. of the IEEE Engg. in Medi-cine and Biology Society (EMBC)Orlando, FL, pp. 1373-1376, doi: 10.1109/EMBC.2016.7590963.

[23] Thao LT, Quang NH, 2017 Automatic skin lesion analysis towards melanoma detection 21st Asia Pacific Symp. on Intel. and Evolutionary Syst. (IES), Hanoi, 2017 pp. 106-111, doi: 10.1109/IESYS.2017.8233570.

[24] Maia LB, Lima A, Pinheiro Pereira RM, Junior GB, Dallyson Sousa de Almeida J and de Paiva AC 2018 Evaluation of melanoma diagnosis using deep features25th Int. conf. on Syst.Signals and Image Processing (IWSSIP), Maribor pp.
1-4, doi: 10.1109/IWSSIP.2018.8439373.

[25] Harangi B, Baran A, Hajdu A 2018 Classification of skin lesions using an ensemble of deep neural networks 40th Annu. Int. Conf. of the IEEE Engg.in Medicine and Biology Soc. (EMBC), Honolulu, HI pp. 2575-2578, doi: 10.1109/EMBC.2018.8512800.

[26] Harangi, Balazs 2018 Skin lesion classification with ensembles of deep convolutional neural networks J. of Biomed. Inform. vol. 86 pp. 25-32.

[27] Mahbod A, Schaefer G, Ellinger I, Ecker R, Ptit A, Wang C 2019 Fusing fine-tuned deep features for skin lesion classification Comput Med Imaging Graph 71:19-29, doi: 10.1016/j.compmedimag.2018.10.007.

[28] ISIC Archive dataset. Available at: https://isic-archive.com/api/v1/.

[29] Diederik P. Kingma and Jimmy Lei Ba. Adam 2014A method for stochastic optimization arXiv:1412.698.

[30] E. Nasr-Esfahani et al. 2016 Melanoma detection by analysis of clinical images using convolutional neural networkX38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, pp. 1373-1376, doi: 10.1109/EMBC.2016.7590963.

[31] Munteanu C, Cooclea S 2009 Spotmole – melanoma control system Available: http://www.spotmole.com/.

[32] Giotis I, Molders N, Land S, Biehl M, Jonkman MF and Petkov N 2015 MED-NODE: A computer-assisted melanoma diagnosis system using non-dermoscopic images Exp. Syst. With Appl Elsevier vol. 42 no. 19, pp. 6578-6585.

[33] Zagrouba E and Barhoumi W 2004 A preliminary approach for the automated recognition of malignant melanoma Image Analysis & Stereology vol. 23, no. 2, pp. 121-135